DOCUMENT RESUME

ED 214 943 TM 820 016

TITLE Options in Education: Takes Listeners to the Core of the Issues. Standardized Testing, Parts 1-6, Program Nos. 280-285.

INSTITUTION George Washington Univ., Washington, D.C. Inst. for Educational Leadership.; National Public Radio, Washington, D.C.

PUB DATE 80
NOTE 90p.

EDRS PRICE MF01/PC04 Plus Postage.
DESCRIPTORS *College Entrance Examinations; Scoring; *Standardized Tests; Student Attitudes; Test Bias; Test Coaching; *Testing Problems; Test Validity
IDENTIFIERS Educational Testing Service; National Public Radio; *Scholastic Aptitude Test; Truth in Testing Legislation

ABSTRACT

The transcript of a six-part National Public Radio broadcast on standardized testing is presented. The first part focuses on the reasons tests are administered; these reasons are discussed by proponents and opponents of testing. Part Two contains a discussion of the possible bias of tests, and their validity. The third part discusses the Scholastic Aptitude Test (SAT), students' experiences in taking the SAT, and students' attitudes toward the scores. Most of the fourth part is concerned with coaching for tests and the implications of coaching for test interpretation. The fifth part discusses the need for more openness in standardized testing, and the proposed test disclosure legislation. The final part of the program criticizes the use of standardized tests and advocates the use of alternative means of student assessment. (BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

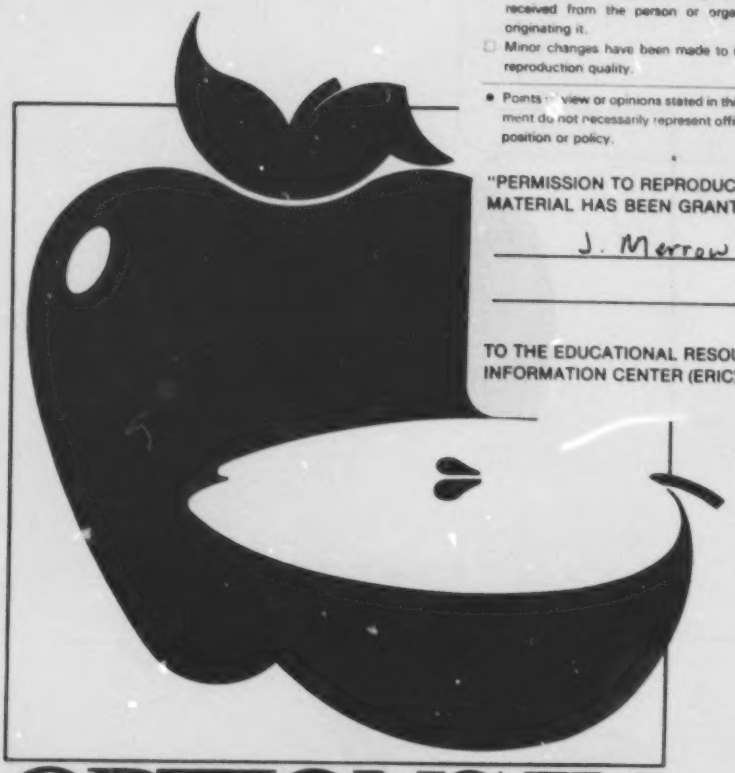
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Merrow

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED214943



OPTIONS IN EDUCATION

takes listeners to the core of the issues

TM820016

STANDARDIZED TESTING, PARTS 1-6, PROGRAM NOS. 280-285

A co-production of
National Public Radio
and the Institute for
Educational Leadership of the
George Washington University

TABLE OF CONTENTS

	PAGE
INTRODUCTION	1
KINDERGARTENERS TAKING THEIR FIRST STANDARDIZED TESTS	1
ACCORDING TO JAMES GUINES THERE'S A STRONG PRESSURE TO RAISE THE TEST SCORES IN HIS DISTRICT, WASHINGTON, D. C.	2-3
A JUNIOR HIGH SCHOOL CLASS BEING TAUGHT HOW TO TAKE STANDARDIZED TESTS	3
EVEN SOME KINDERGARTENERS MUST TAKE STANDARDIZED TESTS. KINDERGARTEN TEACHER ANN HOWE HAS RESERVATIONS AOUT THE PRACTICE	3-4
ROBERT JACKSON OF THE EDUCATIONAL TESTING SERVICE SAYS THAT SCHOOLS GIVE THE TESTS FOR THE WRONG REASONS	4-5
FORMER SCHOOL SUPERINTENDENT FENWICK ENGLISH SAYS THERE NEEDS TO BE A BETTER CORRELATION BETWEEN THE CURRICULUM AND THE TESTS	5-7
JUDAH SCHWARTZ INSISTS THAT IT IS IMMORAL TO RANK-ORDER PEOPLE ON SOMETHING AS COMPLEX AS COGNITIVE ABILITY	7-8
FOR ROBERT EBEL TESTS ARE A WAY OF HOLDING TEACHERS ACCOUNTABLE	8-10
TERRY HERNDON OF THE NEA SAYS TEACHERS HAVE LESS NEED FOR CONCERN ABOUT THE EFFECTS OF TESTING ON THEIR JOBS THAN THE POSSIBLE DAMAGING EFFECTS ON THE CHILDREN	10-12
A CLASS LEARNING TO TAKE TESTS	12-14

"Copyright © 1980 National Public Radio and the Institute for Educational Leadership. This material has been furnished for non profit, non broadcast purposes and maynot be duplicated or otherwise used in violation of the Copyright Law."

STANDARDIZED TESTING, PART ONE

KINDERGARTENERS TAKING THEIR FIRST STANDARDIZED TESTS

TEACHER: Everyone have a point on your pencil. It's very important that we have pencils we can write with. Do you know how to mark? If your teacher did a practice exercise with you, how would you mark it if it were the correct box? What would you do?

STUDENT: Draw a line.

TEACHER: Draw a line. Can someone show me how? You're smart children.

MERROW: Those smart children are only five years old. They're just beginning kindergarten, and already they're taking their first standardized test.

REINHARDT: For the rest of their educational career, they may be judged and categorized according to how they score on tests like these, multiple choice exams with separate, fill-in-the-box answer sheets.

TEACHER: Very good. Now you have to mark it very, very hard and really mark it out, because we're not going to mark these tests at school. They're going out and they're going to be marked by a machine. The machine can't say, "Oh, gee, I'm not sure whether Terry made a mistake or not," can it? So you're going to have to really mark it up so the machine will know don't correct -- don't use this one as the correct answer, but use the other one that you've marked. All right? Any questions?

MERROW: I'm John Merrow.

REINHARDT: I'm Barbara Reinhardt, for NPR's Options in Education, with the first in a series on standardized testing in school.

MERROW: In this program, how the tests and concern about test scores influence what goes on in school in thge curriculum.

REINHARDT: There's really no escaping standardized testing in schools today. This year, 50 million students will take about 200 million of the tests.

MERROW: School systems will pay test makers about $250 million. That's a conservative estimate.

REINHARDT: If you count teacher and administrator time, the cost mounts to well over $1 billion a year.

MERROW: You'll recognize some of the tests. SAT, PSAT, CAT, CTBS.

REINHARDT: The Iowa Test of Basic Skills. The Stanford Achievement Test. The Metropolitan Achievement Test. They're called norm reference tests. Each kid's score is compared to every other score on the same test.

MERROW: Every school system we know gives one or more standardized tests, to find out how students are doing or how the system compares to other systems.

REINHARDT: Critics of testing say scores reflect parental income and social class and shouldn't be used to judge schools, teachers or students.

MERROW: But if your school system scores are low, that stance is a luxury you can't afford. According to James Guines, there's strong pressure to raise the scores in his school district, Washington, D. C.

JAMES GUINES SAYS THERE'S A STRONG PRESSURE TO RAISE THE TEST SCORES IN HIS DISTRICT, WASHINGTON, D. C.

GUINES: The pressures come from the community and from the Board of Education. It comes from the hierarchy itself. In other words a Superintendent is measured, for some reason, by how well students do on standardized tests. He holds the lower ones of us accountable for what contribution we can make to making those things happen. My concern, honestly, about that, is I've never been fully convinced at all, as an educator, that standardized tests, in and of themselves, represent what we would want to believe to be the educated person. The whole concern for feelings, appreciations of beauty, appreciations of art, recreation, interpersonal relations, I'm just not sure that standardized tests in any way address those kinds of broader dimensions of what we'd like to see in the educated person.

MERROW: There's a conflict.

GUINES: Not a conflict. A question. A question. As a public school paid administrator, I certainly, and the Superintendent must, be somewhat responsive to the demands made upon them by the community, in terms of their expression of what they consider to be good schools. We have to be responsive to that. We have to do that to get over at least the credibility situation. And I guess what I'm saying is let's get on with that. Let's get that out of the way and then commence to address some of the larger, larger issues and work with our community, to begin to understand some of those larger, larger issues, about what constitutes an

educated person.

MERROW: Then shouldn't you be educating the community about the relative importance of standardized tests? That is to say that they're not all that important?

GUINES: We really do. But when you are the lowest on the totem pole, sometimes that's thrown back at you as copout.

REINHARDT: James Guines, Associate Superintendent for Instruction in the Washington, D. C. Public Schools.

MERROW: Washington tests its kids often, beginning in kindergarten. Some public schools even have courses in how to take standardized tests.

A JUNIOR HIGH SCHOOL CLASS BEING TAUGHT HOW TO
TAKE STANDARDIZED TESTS

TEACHER: Our objective today is to review the use of context clues. Now, who can remember what we said that context clues were? What is a context clue? Anybody? Leah?

LEAH: A context clue is when you have a word or something and you don't know what it is, but something in the sentence will tell you what the word means.

TEACHER: Very good. Okay.

MERROW: In Washington and in other school districts as well, test scores are so important that schools teach test taking skills.

REINHARDT: But that's a small part of the curriculum. We're wondering how the tests and the test results are connected to the rest of the academic program.

MERROW: One would think teachers could use results to determine where students need help. Not always possible, says kindergarten teacher Ann Howe.

EVEN SOME KINDERGARTENERS MUST TAKE STANDARDIZED TESTS
KINDERGARTEN TEACHER ANN HOWE HAS RESERVATIONS ABOUT THE PRACTICE

HOWE: When the results come back on these tests it's so late in the year that it really doesn't make much difference.

GOLDFIELD: When do they come back?

HOWE: Oh, I probably won't get these results for another month or so.

GOLDFIELD: And what are they used for when you get them?

HOWE: When it comes back, I go through it and see if it really applies to what I'm doing with the children, and it just does, and most of the time you've passed everything in there anyway. You know, if it's rhyming or just visual perception or whatever it is. Unfortunately, it's just one of those things that's rather late, and it's a very basic test.

GOLDFIELD: Is it ever used for placement purposes?

HOWE: Not by me, and I don't know that it could be by most people, because by the time you get the test results back, you must surely have done whatever placement you want to do.

However, I don't group children in any kind of rigid group, so it doesn't make any difference. But -- excuse me, hon. Excuse me. I'll talk to you in a few minutes. I know. You'll do your test tomorrow, okay? It is a method of showing parents how well they did. I think that's really what it's used a lot for, so that when you have conferences with the parents you can say, "See, this is how your child tested," and, "This is what I needed to work with your child," or you know, that sort of thing.

If a child is still low -- suppose that a test comes back and you are finding that the child is low in someting and he tested low and is still, well, you can use the test when you're talking to the parents and say, "See, he tested low in this and this is a problem for him and perhaps he does have a visual problem," or you know, "Perhaps he has a hearing problem or something." So sometimes you use it that way.

MERROW: Kindergarten teacher Ann Howe. Rebecca Goldfield talked to her at her class at Stevens School in Washington, D. C. That's a common criticism, that standardized test results come back too late to be useful in improving the curriculum.

REINHARDT: But the test makers say the real problem is not delay. They say schools don't plan. They give tests for the wrong reasons. Robert Jackson of Educational Testing Service.

ROBERT JACKSON OF THE EDUCATIONAL TESTING SERVICE SAYS
THAT SCHOOLS GIVE THE TESTS FOR THE WRONG REASONS

JACKSON: Often, school districts don't look at tests and testing as something that should support the instructional process, but look at it as a necessary evil that they have to go along with.

MERROW: You're talking about the larger political swirl that goes around the school.

JACKSON: Precisely. Precisely. And it's often in that context that tests are introduced into the system. Oh, there are notable exceptions, I'm sure, but the point is that the concept of -- a living concept of testing to support instruction, I think by and large is absent in the school systems throughout this country.

HERROW: A living concept of tests to support instruction? It sounds mystical.

JACKSON: I don't think it's at all mystical. It's just simply my way of saying ethical behavior.

HERROW: I feel that youngsters tend to be the victims in educational enterprise. In the situation I described in which there is no harmony between the behaviors being taught and the behaviors being measured, in those situations the students are the ones being labeled. Sometimes it's so mudane and callous as to be called dummies. Other times they're culturally deprived or impoverished or whatever.

The real situation might be that if we had better harmony between what we're teaching and what we're testing that we'd get a greater reflection of student performance.

HERROW: Okay. Well, let me argue that suppose you have a system which spends three or four days effectively using instructional time to give a norm reference test to its third graders, doesn't use the results to any way change the behaviors of those third graders, doesn't take the test and go over it with the kids. My argument would be that that's wasted time. All you get then is nu mber which says you are in the 82nd percentile, or the 22nd percentile.

JACKSON: In that scenario I would go further than to say it's just a waste of time. I'd say it's unethical behavior.

REINHARDT: Robert Jackson of Educational Testing Service in Princeton, New Jersey.

HERROW: Former School Superintendent Fenwick English says that if school systems knew what they wanted to teach they could use the scores wisely.

FORMER SCHOOL SUPERINTENDENT FENWICK ENGLISH SAYS THERE NEEDS TO BE A BETTER CORRELATION BETWEEN THE CURRICULUM AND THE TESTS

ENGLISH: The first thing you look for is what are the things that you want kids to really learn, and that should be specified to the extent that it's accessible, either through standardized tests, which are called norm referenced or criterion referenced, or through

other means, such as teacher observation, other types of assessments which aren't the multiple choice type computer processed assessment. There are many ways you can assess children and what they learn in a classroom, but they ought to be related to what we think children ought to know.

And we haven't really done -- I say we -- the profession has not done a very good job, nationally, on specifying what it is kids ought to learn. So the first thing you do is you find out what it is children should learn. And then you go about defining your assessment program accordingly.

But what's happened is we have these real mushy goals that we -- which we call motherhood, blue sky, apple pie and the flag, which no-one can argue with but no-one really knows what they are. There's a thousand different ways to translate them. And so when you get the test data in which is very specific and hones in on you know, just a few specific examples, then you say, well, that's not what I meant by citizenship, or no, no, no, no, we have another way that we -- so therefore we didn't determine that the tests matched what our objectives were initially, going into this thing. And then of course when your scores look bad it looks like you're giving a pile of excuses about why, you know, the kids have not done well, when really, you didn't have a match between what you thought the children ought to learn, and the test that you selected or was selected for you, to assess their learning.

REINHARDT: And that's being generous. In many cases you didn't really even decide exactly what the kids ought to be learning.

ENGLISH: That's right. In many cases that was almost left wide open. And the test comes in. You cannot write a test. The test is specific. It's going to look at specific things. And in many cases that's why the test has defined the curriculum, and it exerts a control on the curriculum that many teachers, I think rightly, resist, because they -- if you understand how tests are made, they're made by human beings and they're made by people and there are a lot of biases involved in the development of tests and so -- and they've been misused, I think, in that respect. I don't think any school district can have local control if it doesn't really know what the test is assessing in the first place that it selected. It may be assessing somebody else's curriculum, but not the local district's curriculum.

REINHARDT: Are you suggesting that each school system has to be so specific in each grade as to decide exactly what skills, the use of commas in the fifth grade, you know, et cetera, et cetera, all of the skills, specifically in long, long lists?

ENGLISH: Absolutely. And it's not that difficult. But unless

you've done that, you can't use a test constructively or positively. Believe me, the test makers have to do that. So in other words, if you leave your curriculum open ended for whatever reason, and they range fropm not wanting to impinge on somebody's creativity to just sheer laziness, don't want to really do it, you can't write a test that way. The test is written and it's based on a fact that there are assumptions made about what kids ought to learn by a grade.

So if you leave your curriculum open ended and you use a test that already has a whole bunch of assumptions made about what children are learning, you're adopting the test maker's curriculum, and there's no other way to say it.

MERROW: Fenwick English, former School Superintendent, now partner in the Washington consulting firm of Pete, Marwick and Mitchell.

REINHARDT: What should be the connection between standardized test scores and what goes on in school? Judah Schwartz says there should be no relationship.

JUDAH SCHWARTZ INSISTS THAT IT IS IMMORAL TO RANK-ORDER PEOPLE ON SOMETHING AS COMPLEX AS COGNITIVE ABILITY

SCHWARTZ: I believe it is deeply wrong to rank-order people and to adopt the stance that one has the technology for collapsing the complexity of individuals into a single number and arraying those numbers on a line. I believe it borders on immorality. Under no circumstances should people be rank ordered, even assuming that it could be done. They can, of course, be rank-ordered for such properties as height and weight, but they cannot be rank-ordered for any multi-dimensional trait, despite the mathematical pretentions of my colleagues in the social sciences. It's a fundamental absurdity. You cannot rank order anything unless it's one dimensional, and I defy anyone to say to me with a straight face that anything like cognitive ability in any domain is one dimensional or can be aggregated into something that is one dimensional.

You can ask a child to multiply and then you can look -- give him a multiplication problem and then you can look at a very large number of other children and rank the child as to how well this child multiplies compared to the other children. And in fact you can do that with a large enough number of items and then discover that your child multiplies better than 73 percent of the children in Louisville, Kentucky, which might have been the norming population. But it's of very little concern to me that my child multiplies better than 73 percent of the children in Louisville, Kentucky, if he doesn't know how to multiply. And if he does know

how to multiply, then I don't care whether he multiplies better or worse than any percent of the children in Louisville, Kentucky.

ROUSE: How can tests diagnose a child's ability in something?

SCHWARTZ: : Tests can't diagnose anything unless they're built on a deep understanding of the structure, the cognitive structure, of the task. No test to my knowledge right now that is available to the schools is built on any kind of sophisticated understanding of the structure of any of the tests in either language or mathematics.

You see, there's a fundamental misunderstanding in all of this, and that is that we assume in using these tests that if a child performs that he's competent, and if a child doesn't perform that he's not competent. That's an absurdity. It's also logically inconsistent. Performance may imply competence, but non-performance does not imply anything.

ROUSE: Is there one form of testing that you would advocate over another, or are you in favor of abolishing both kinds?

SCHWARTZ: I'm not at all in favor of abolishing tests. I think it's absurd. The society has a real need for information to inform instruction. It has a real need for information for accountability purposes. It has a real need to select for access to limited resources. We hae to have procedures. I reject present procedures, which I think are intellectually fraudulent and ill-grounded and do an immense amount of disservice to various and quite significant fractions of our population.

REINHARDT: Judah Schwartz, a Professor at MIT, talking with Mary Rouse.

MERROW: But even if a single score of one student may be misleading at times, reporters of standardized tests say tests are the quickest, cheapest and most reliable way of holding schools accountable.

REINHARDT: Or of holding teachers accountable. Professor Robert Ebel believes that schools should give more standardized tests and then use the results in a more rational, more precise way.

FOR ROBERT EBEL TESTS ARE A WAY OF HOLDING TEACHERS ACCOUNTABLE

EBEL: Any teacher ought to be responsible each term for producing some evidence that the teaching has done some good, namely that the students have learned something.

MERROW: Aha, so it's the norm referenced tests are not just being given to judge the students. They're being given to judge the teacher.

EBEL: Yes. From the teacher's point of view, they're being used to judge the effectiveness of the students' efforts to learn, but from an external point of view, they also can indicate whether the teacher has been doing a good job of teaching.

MERROW: So let me design a school system that I think you would buy into. At the beginning of the year we'll give standardized tests to the kids in, oh, mathematics. At the end of the year we'll test those same kids, and if they haven't learned then we're going to draw some conclusions about the teacher?

EBEL: I think that's correct. I'd want the teacher to have a day in court, to be able to explain if they haven't learned why they haven't learned. It certainly may not be the teacher's fault, but the aim of the school is to help them to learn, and we ought to be assessing how well that's happening.

MERROW: How about the principal? Will you hold the principal responsible?

EBEL: To some extent. I think the teacher, having done this job of trying to provide evidence, should sit down with the principal and say, "Now, here are the problems I have and there are some of these that I think you can do something about." And so I think the community should hold the principal partly responsible for maintaining a good learning environment.

MERROW: What happens to the teacher whose students do not learn?

EBEL: I would say that if it is the judgment of the principal and the supervisor that this is because the teacher is an inept teacher or poorly trained teacher, then the teacher ought to be encouraged to find other work.

MERROW: Oho, very cut and dried like that.

EBEL: Well, the decision isn't going to be cut and dried. It's going to be a very difficult decision. But you know, football coaches get fired if they don't win. And I don't see any reason why teachers ought not to be encouraged out of the system if they can't succeed in encouraging learning on the part of their students, or facilitating that learning.

MERROW: Let me put forth this proposition and get your reaction to it. No school, no teacher, should ever give a norm reference test unless it were immediately possible to go over -- in fact required, to go over that test with the people who just took it, to discuss what's right, what's wrong and how the reasoning might have worked to get to a certain answer, unless there's an

immediate feedback, I guess is the word.

EBEL: Yes. I would reject that proposition.

HERROW: Why?

EBEL: It seems to me that the information to the teacher and to the student should indicate, A., things are going well, B., things could be going better, C., things are going very badly. And we need to do something about this evidence of deficiency or success in -- we don't need to do anything if it's successful, but this is information we can use to guide our future efforts. Not in minute details about why did you answer this question wrongly. The kind of questions that appear on a standardized test are like those that have appeared in profusion in the textbook and in the work manuals and in the work books. So it isn't necessary that you consider the standardized test as a sort of curriculum that has to be mastered. Instead, you consider it as an indication, an overall summary indication of how well things are going. And if they're not going well, you go back and try to do them better.

HERROW: Robert Ebel, Professor in the College of Education at Michigan State University, and a staunch defender of standardized testing.

HERROW: We asked Terry Herndon of the National Education Association whether teachers are afraid that student standardized test scores will be used for hiring and firing.

TERRY HERNDON OF THE NEA SAYS TEACHERS HAVE LESS NEED FOR CONCERN ABOUT THE EFFECTS OF TESTING ON THEIR JOBS THAN THE POSSIBLE DAMAGING EFFECTS ON THE CHILDREN

HERNDON: Well, there are constantly people at work in the political system, advocating that teachers be evaluated that way. So there is of course some teacher anxiety about that. But when I'm dealing with teachers and they raise that as their anxiety, I simply pose the question, "Is there a teacher in the room that has ever been adversely affected by that assumption?" And normally there are none. And my response is that then I think we ought to worry about that when it happens. Right now there are kids being adversely affected by these tests.

So if and when there are efforts, and there have been a few to evaluate teachers on the basis of students' performance on the standardized tests, we will fight it. But we are not -- we're not preoccupied with that dynamic.

HERROW: You're saying now that what you are concerned about is kids. You are in effect kind of wrapping yourself in the flag.

But you'll have to allow me to be a bit skeptical. You are known for being able to look down the road, and that's what you're doing:

HERNDON: I understand the skepticism. When I say kids I'm not only talking about five, eight, twelve, fifteen year old children. I'm also talking about graduates that are entering college. I'm talking about graduates coming out of college, applying for a teaching certificate. I'm talking about a lot of categories of people that NEA relates to.

MERROW: Let's talk about the use teachers make of norm referenced tests and the results. Let's just conjure up a picture of a school year in which, say, at two or three points during the year the kids will take some kind of standardized, machine scored, norm referenced or criterion referenced test. What happens then? Does the teacher make use of that test, the test results, in some way?

HERNDON: If I cited one very prominent school system that does a lot of testing, it happens to be the one where my children go to school, I've asked the school administration what they do with all of those tests. They tell me they don't use them for placement purposes. I say, "Well, then, what do you do?" They say, "Well, we use them to help fashion instructional programs for individual students." I say, "Is it not true that all the teacher gets in terms of my daughter's test results is the raw score and a percentile rank?" The answer is yes. How is that used for instructional planning purposes? The answer is exceedingly vague.

In most cases I think the tests are used particularly in an affluent suburban school district, to prove what everybody already knows. The children in that school district are achieving at above national norms, and that's supposed to persuade the community that the school district is doing an outstanding job.

MERROW: The purpose of the test is political in a large sense, rather than instructional or school related.

HERNDON: I think 90 percent of the time that's so. I've indicated in other fora, John, that there are three reasons for making decisions in schools. One is political. One is administrative. And the third is pedagogical. And in most cases the tests are used for political purposes. The second most prevalent motive is administrative. And down at the bottom of the line there are a handful of pedagogical decisions that can use these data.

MERROW: Someone from ETS, the Educational Testing Service, just told me that in most classrooms the tests are unrelated to the curriculum.

HERNDON: I'm not sure of the import of that statement. If you're talking about standardized tests, when they first enter a school system that's generally true. The test is designed by people that play no role in the evolution of the curriculum in that school district. But once the test scores themselves become politicized, the school system adapts very readily and begins to fashion a curriculum that will lead to higher test scores and therefore a higher level of political satisfaction in that community.

MERROW: But that's okay, if the test is valid.

HERNDON: Well, it's okay if the school board, the teachers, the parents and others have analyzed that test and have made a judgment that the test really does reflect the curriculum we'd like to see in our school system. But if there is an assumption that high test scores have intrinsic value and therefore, without examination of the curriculum that is associated with the test, we will fashion a curriculum that assures high scores, I don't think that's valid at all if you have any confidence at all in a locally based decision about what a school ought to be doing.

REINHARDT: Terry Herndon, Executive Director of the National Education Association in Washington, D. C.

MERROW: The relationship between standardized tests and what is taught in school is a confusing one, often because it's not clear which is more important.

A CLASS LEARNING TO TAKE TESTS

TEACHER: For our activity we're going to do the following sentences. Now, I'm going to call on you and you tell me what words in the sentence would allow you to know the meaning of the word that you are supposed to be unfamiliar with.

All right, let's take number one. All right, Shawn. Not Shawn, Gerell.

GERELL: Charades is a game in which players must guess the meaning of a word from watching other players act out a dramatic representation of it.

TEACHER: Okay. Charades is a game in which players must guess the meaning of a word from watching other players act out a dramatic representation of it.

What word or words in that sentence that you know, a familiar word, which is going to be charades? Which words? James?

JAMES: Dramatic representation.

TEACHER: Okay. That's some of the words. Allan?

ALLAN: Watching the other players?

TEACHER: That's some of the words. Roger?

ROGER: Games.

TEACHER: Games. That's some of the words. Deborah?

DEBORAH: Players.

TEACHER: Players. But here you need most of those words to let you know what charades means. Okay? Because the whole sentence gives you the defi--what? --nition.

STUDENTS: Definition.

TEACHER: Definition. It's a game. That's true, too. But again you have definition and description of your unfamiliar word, charades. Now, do you know what context clues are? Are you getting an idea of how to find definitions or meanings --

HERROW: Teaching kids how to take standardized tests, at a junior high school in Washington, D. C.

REINHARDT: The system is assuming that test scores are meaningful, that standardized tests do measure what we want children to learn. But what if the tests are biased? If the tests really measure something else?

HERROW: We'll take up those questions in the second part of this series. I hope you can join us.

TEACHER: Everybody gets one of these.

REINHARDT: Now what you're doing is passing out standardized test forms.

TEACHER: Right, because we're getting them familiar with gridding, how gridding the whole thing, covering the whole thing, shade it. Sometimes, like we had one activity in which I told them to don't shade in the whole circle or square, to just put a little spot there to save time, because all our standardized tests are timed, and come back when they've finished and shade it all in.

STUDENT: I need a number two pencil.

TEACHER: No, you don't need a number two pencil. I'm going to check these out myself.

STUDENT: So we can use a pen?

TEACHER: Yes.

REINHARDT: So they're even used to using number two pencils.

TEACHER: Right, because when they see something like this, they know they're supposed to use a number two pencil.

REINHARDT: A complete printed transcript of all six parts of this series is available for $6.00. The set of cassette recordings costs $18.00.

MERROW: Write to us at National Public Radio - Education, Washington, D. C. 20036.

REINHARDT: $6.00 for the printed transcript of all six programs. $18.00 for the set of cassette recordings.

MERROW: And our address again, National Public Radio - Education, Washington, D. C. 20036.

REINHARDT: Options in Education is a co-production of National Public Radio and the Institute for Educational Leadership of the George Washington University.

MERROW: Support for the series is provided by the National Institute of Education and the Corporation for Public Broadcasting.

REINHARDT: John Merrow is the producer of Options in Education. Mary Rowse is our Associate Producer and Rebecca Goldfield is our Assistant Producer.

MERROW: Barbara Reinhardt Serves as co-host and reporter, and we've had engineering assistance for this program from Mary Ishnaga and John Widoff. I'm John Merrow.

REINHARDT: I'm Barbara Reinhardt.

MERROW: This is NPR, National Public Radio.

Program No. 281: Standardized Testing, Part Two

TABLE OF CONTENTS

	PAGE
INTRODUCTION	1
JUDAH SCHWARTZ CITES A TEST QUESTION HE BELIEVES IS MALICIOUS	1
GERDA STEELE EXPLAINS WHY BLACK STUDENTS DO NOT PERFORM WELL ON THE TESTS	1-2
CHUCK STONE, FORMERLY OF ETS, SAYS THE TESTS MEASURE INCOME MORE THAN ANYTHING ELSE	3-4
ALAN NAIRN, WHO AUTHORED A REPORT ON STANDARDIZED TESTING, CONCURS	4-5
JOHN FRIMMER OF ETS SAYS THE TEST MIRRORS REALITY	5-6
WILLIAM TURNBULL THINKS THE TEST IS EXTREMELY VALUABLE	6-7
CAROL DWYER SAYS THE TEST PROTECTS WOMEN	7-8
JOHN WEISS SAYS THE TEST FAVORS THOSE WHO THINK QUICKLY AND HURTS THOSE WHO THINK METHODICALLY AND CAREFULLY	9-10
BANESH HOFFMAN SAYS THE TEST DISCRIMINATES AGAINST THE MOST BRIGHT AND CREATIVE STUDENTS	10-12

"Copyright © 1980 National Public Radio and the Institute for Educational Leadership. This material has been furnished for non profit, non broadcast purposes and maynot be duplicated or otherwise used in violation of the Copyright Law."

STANDARDIZED TESTING, PART TWO

JUDAH SCHWARTZ CITES A TEST QUESTION HE BELIEVES IS MALICIOUS

SCHWARTZ: There is an example of a problem taken from the Metropolitan Achievement Test, in which a child is offered a picture of a triangle, a rectangle, a circle and a hexagon and asked which figure has the largest number of sides. Now, if you ask me what the right answer to the question is, I do not know. It is possible to have two quite different cognitive representations of a circle. One is as the limit of a polygon with an infinite number of sides, and one is as a continuous curve.

The problem is deceitful, malicious and built on ignorance.

REINHARDT: A child who thinks creatively might well have trouble on that question.

MERROW: So might a little girl who's been taught that math is for boys.

REINHARDT: In fact, according to critics, many different kinds of people perform poorly on standardized tests, regardless of intelligence. It's not clear why they have problems with these machine scored, multiple choice exams, but their low scores may seriously limit their future.

MERROW: Are low scores the fault of these groups or of the tests themselves? I'm John Merrow.

REINHARDT: I'm Barbara Reinhardt, for NPR's Options in Education, with the second part of our report on standardized tests. This time, the question of bias.

MERROW: According to Gerda Steele, racial minorities get hit the hardest.

GERDA STEELE EXPLAINS WHY BLACK STUDENTS DO NOT PERFORM WELL ON THE TESTS

STEELE: When you think of atrocities you think of death, you think of dying, and I believe that the way testing has been used in this country, it has been an atrocity committed against minority students, because a lot of students have died, not physically, but mentally and emotionally, because of the effects of testing on them.

Let's say you have a child that's in a situation in a school where that child knows and feels and believes that the teachers, they don't think that he can learn, that he's bright.

They think he's stupid, he has a low IQ. Guess what happens? It becomes a self-fulfilling prophesy. The child actually becomes that. And it has a lot to do with, like I said, expectation. If you expect something from a child and you try to bring them to a certain point and move them along, that child will work with you. But if that child comes to you and that child sees that you don't feel there's any hope for him or her, then they just give up.

MERROW: And you're saying that the kid gets a bad score on a test, the teacher learns the score and says, "Oh, she's not very bright. He's not very bright."

STEELE: Right. Then also, as a part of that, then that student is trapped into -- it could be a mentally retarded class or a compensatory class. And not the fact that they're trapped into that just maybe for one semester or one year. A lot of kids are trapped into that for the rest of their lives.

MERROW: But black kids don't do as well on tests as white kids do.

STEELE: Well, because most tests are normed on white children. 99 percent of your test makers are whites. One of the assumptions that a test makes is that when kids come to a test they come with similar experiences, learning experiences, similar educational experiences. And since you have so many unequal opportunities in education, that is not true. So if you are of the majority population and you have gotten everything that you're supposed to get in terms of an education, and if you are of a minority population and have been denied certain kinds of opportunities, how can the minority do as well as the majority in that particular situation?

MERROW: So that what's unfair is not the test, it's educational opportunity.

STEELE: It's both. It's both. Because this is what happens. the educational opportunity is unfair, all right? Then the test is unfair because the test assumes that that child has gotten everything they were supposed to get out of that educational experience, and that does not happen.

MERROW: Well, maybe that's not what the test is doing. Maybe the test is saying, well, let's see who's gotten what.

STEELE: Okay. What tests really do is sort people out. That's what they really do. They classify. Okay, you're a green, you're a bluer, you're a yellow, right? And one of the things with that is that gives you a label that may stay with you for the rest of your life.

MERROW: Gerda Steele, former Education Director of the NAACP.

REINHARDT: On one standardized college admissions test on which the maximum score was 1,600, blacks as a group scored 255 points lower than whites.

MERROW: For Chuck Stone, who used to work for one of the test making companies, ETS, the issue is not so much race as class.

CHUCK STONE, FORMERLY OF ETS, SAYS THE TESTS MEASURE INCOME MORE THAN ANYTHING ELSE

STONE: Are standardized tests are normed on middle class populations, and the middle class populations in America are white. For example, in the median family income, there's a discrepancy of roughly $8,000.00 between black and white kids. So that ETS will even indicate in its tables that there was a high correlation between test scores and income, so that at $18,000.00 your test scores of kids would be, oh, about four hundred and something, and at $10,000.00, which is what -- where the average black kid is, the test scores are about 350 or 300. So there's this discrepancy in the test scores, which is a function of the socio-economic status.

MERROW: You seem to be saying that you don't really need the tests. All you need is the data about income and so on?

STONE: I'm saying -- what I'm saying is this. You show me any two random distribution of -- two sets of random distribution -- randomly distributed test scores, and I will tell you which ones are from Princeton and which ones are from Newark, which ones are black which ones are white, which ones are wealthy, which ones are poor. Those are the factors that shape the test scores. Neighborhood density is a factor in the test score result.

Your whole life gets distilled into the funnel of a three hour test and you can be wiped out. And the other thing it does, and this is what Warner Slack and Doug Porter, in this brilliant Harvard Education Review pointed out, that it demolishes the pupil's or the test taker's self-esteem. You might be a good student doing very well and you're motivated to achieve. You take the SAT -- in fact it almost hurt my daughter. That SAT score hurt her because she did not do well on the SAT's, and she went to the coaching academy and still didn't do well. But she did extremely well in school. Physics, math. Got A's in those subjects.

MERROW: What did you say to her?

STONE: I said -- of course I gave her the standard argument, tests are a bunch of nonsense. They don't do a thing. They are not an accurate predicter. And I showed her the tables, the

correlation, the validity coefficients, and showed her that they didn't predict as accurately for lower test scores as they do for higher test scores, and she was okay. But I had to shore up her self-conception. It was almost destroyed when she got her test -- I mean she was really hurt.

REINHARDT: Chuck Stone, former Director of Minority Affairs at Educational Testing Service. He's now an editor with the Philadelphia Daily News.

ALAN NAIRN, WHO AUTHORED A REPORT ON STANDARDIZED TESTING, CONCURS

NAIRN: It's not just a difference between rich and poor. This is an important point. Students from wealthy families score higher than students from upper middle class families, who score higher than those from middle class families, who score higher than those from working class families, all the way up and down the line.

REINHARDT: Alan Nairn.

NAIRN: At the same time that these tests are ranking people according to their social class, they have little connection to how well you will actually do when you go into school or when you go into your later career. I mean it doesn't just measure social class. There are other factors. Like, for example, the specialized skill of test taking, which is a kind of novel ability that some people can learn. There are now many schools that charge a great deal of money to teach you that skill, which ironically tends to reinforce the class distinction.

MERROW: The test measures how well you did on the test?

NAIRN: Basically, yes. I mean if you're talking about it in strictly objective terms, that's really all it measures.

MERROW: But that indicates social class, parental income.

NAIRN: Right, because the tests tend to be class bound in terms of the kinds of questions they ask, the kind of vocabulary they test, and so on.

MERROW: But they don't ask those questions that people talk about where you are asked a question that requires you know how to play croquet. I mean they've eliminated questions like that.

NAIRN: They';ve eliminated the more blatant ones, but the pattern of distribution of scores by social class is still there. Now the biases are more subtle.

MERROW: ETS in response says these tests do not create distinctions, they merely reveal them.

NAIRN: Right. ETS claims that they are just an impartial mirror that is showing the inequalities in society. But that's simply not true, because if you look at how people perform once they get into college, what kind of grades they actually earn, you find that the grades people earn are not related to their social class in the way that the test scores are. Ranking by test -- by SAT scores, say, is a ranking by class, but a ranking of first year grades is not a ranking by class. There was a huge study done by the American Council of Education involving more than 50 colleges and 30,000 students, which came to that conclusion. Even though your social class tends to determine where you'll end up on the SAT, once you get into college, that washes out.

ETS itself did a study last year, which showed that when you look at outside accomplishments, outside the classroom, and things like writing, scientific activities, leadership, et cetera, when you look at those measures of actual accomplishment, there's no correlation with social class. So this was a judgment of people on merit, on accomplishment, and it had nothing to do with social class. Of all the indicators of supposed merit you can use, the test score is by far the most class biased.

REINHARDT: Alan Nairn, the principal author of, "The Reign of ETS, the Corporation that Makes Up Minds."

MERROW: ETS and other test makers reject the charge that their tests are biased against any social or economic group. Dr. John Frimmer of ETS told a college group recently that admissions tests like the SAT do a service by measuring academic skills.

JOHN FRIMMER OF ETS SAYS THE TEST MERELY MIRRORS REALITY

FRIMMER: Just think about this. Reading skills are important in college, right? Creativity is important. Motivation is important. Stamina is important. But so is reading. What are you going to do if you can't read? And that's one of the things we measure with our tests. And reading skills are reading skills whether you're rich or poor. It's not reading skills for the rich and reading skills for the poor.

That's interesting. Well, I may hear more about that. Just to respond to that low roar, though, I believe, and I think I could back it up with time, that if you come from a poor family and you have good reading skills and you take an admissions test, you'll get a high score. If you come from a rich family and you have lo reading skills and you take an admissions test, you'll get a low score. And what the admissions test tells you is what

reading skills you've developed.

Now, it's true that the more money you have the better the chance you have to develop those skills. But I don't think that's a flaw in the test. That's just in the nature of our society. Affluence leads to opportunity and tests reveal the differences.

REINHARDT: John Frimmer, of ETS, speaking at William Patterson College in New Jersey.

MERROW: The test maker's defense goes beyond arguing that test results mirror reality. According to ETS President William Turnbull, it would be irrational and irresponsible not to use such tests as the SAT.

WILLIAM TURNBULL THINKS THE TEST IS EXTREMELY VALUABLE

TURNBULL: There's a few things in the world that are fairly -- whose meaning is fairly clear. The SAT score has the advantage that a 400 or a 500 has meant the same thing year after year. It is a stable currency in a fluctuating world.

The problem with high school grades, for example, is it's so difficult to know what a grade of B or C means. It means different things from different schools. That's not to say that those grades are unimportant. They're quite important. But something like a test score that does mean everybody has run the same race, it has the same meaning from year to year, it does not suffer from grade inflation, perhaps takes on in people's minds a greater solidity, and --

MERROW: Greater than it deserves, or greater than other things?

TURNBULL: Greater than it deserves for some people, although I think that there is merit in having a kind of standard currency that is not devalued over time.

For example, are you going to admit a student to a class in calculus. I don't think of it as a goody, to admit a student who has not learned basic algebra, to admit that student to calculus. I think that what you should say to a student is, "Unfortunately you have not yet prepared yourself to the extent you should have before you can benefit fully from this course in calculus, and then the student should not feel deprived of the opportunity to learn calculus, but well advised to go back and fill in his or her background to the point where the substance of that course will be relevant to the degree of preparation that has been achieved.

The SAT measures the kind of reasoning ability and the degree of reasoning ability the student has developed at a particular point in time.

MERROW: Of course it's not only the student who reacts to the score. It's also the dozen or so colleges that he or she's applied to.

NAIRN: Yes. If you take two students, one with a high SAT score and the other with a low SAT score, you may say it's the SAT's fault that the one with the low score is not being admitted to college. If, however, the college were to admit blindly and take that student, rejecting the one who is prepared, I would say the college has an equally grave responsibility in turning away the student who had worked hard, who had developed his ability or her ability to a high point and was now being turned away from an opportunity in favor of a student who had not. And I don't see that society would be the gainer from that exchange.

REINHARDT: William Turnbull, President of Educational Testing Service, defending standardized tests. But the charges of cultural bias, racism and sexism endure.

MERROW: Critics are concerned about test questions that make assumptions about sex roles.

REINHARDT: A math question in which a girl has to figure out how many eggs to use if she doubles the cake recipe.

MERROW: Carol Dwyer writes test questions for Educational Testing Service. She says that test makers have become so aware of potential sex bias that it doesn't have a chance of getting into a test today.

CAROL DWYER SAYS THE TEST PROTECTS WOMEN

DWYER: In many kinds of educational materials, in all kinds of educational materials up until the recent past, we have seen women treated in a number of ways that we don't see now. I'm talking about textbooks and teacher preparation materials, as well as tests. We have taken the point of view at ETS for a long time that any material that's potentially offensive to women or to minorities doesn't belong in a test. And we don't need to have demonstrated to us even that it's going to make somebody have a lower score. If it's there it's offensive, it's potentially upsetting, it has no place in a test or in any other educational material.

Separate issue from what kind of score you're going to get. It's inherently unfair to have that kind of material in a

test, so we don't.

MERROW: The fundamental point you want to make is that standardized tests are important to women. Tell me.

DWYER: Standardized tests are -- one of the ways that women can establish their credentials. It gives you something to point to when you're in a situation where people are making judgments about you. For college or graduate school admissions, for jobs, for licensing and certification, it gives you the kind of thing that a woman can point to and say this represents my achievement and no-one can question it. We don't get into the realm of things in other situations that people would call intangibles, whether in an interview, for example, somebody says does this person show promise or not, or somebody says this person doesn't really seem to be on the ball. You know. Here is something that a woman can point to that's outside of her personal characteristics. It's outside of the realm of judgments that people can make about her, and it says this stands alone to represent my achievement.

MERROW: Outside the old boy network, is what you're really talking about.

DWYER: That's right. That's right. When I see people talking about abolishing admissions testing, that drives me crazy. I think about where the woman is left then in the selection situation. I mean we talk about back to truth in testing. I look at the people who are advocating that and I say, "I don't see women well represented there. I don't see anyone worried about what's going to happen to women when that test is not there." You know, I see back to sexism.

REINHARDT: Carol Dwyer of Educational Testing Service.

MERROW: When you're taking a standardized test, you have only so much time to finish. That may discriminate against students who work methodically and slowly, including some very bright ones.

REINHARDT: Another group poorly served by standardized tests.

MERROW: Test makers admit they can't easily measure intellectual depth, so they use speed as a surrogate, a stand-in.

REINHARDT: For the test taker, the weight given quick thinking, along with the test maker's penchant for tricky questions, can cause problems. John Weiss.

JOHN WEISS SAYS THE TEST FAVORS THSE WHO THINK QUICKLY
AND HURTS THOSE WHO THINK METHODICALLY AND CAREFULLY

WEISS: One of the things I've found out, and I've given this test to three different groups of people who were social scientists, humanities professors and physicists, and the questions were the following. The first question was spot, spell spot.

MERROW: S-p-o-t.

WEISS: What do you do when you come to a green light?

MERROW: Stop.

WEISS: You go when you come to a green light, but you got tricked by the question because you melded it together. Take another question, for instance. Spell folk.

MERROW: F-o-l-k.

WEISS: Spell Polk.

MERROW: P-o-l-k.

WEISS: What's the white of the egg called?

MERROW: Yolk.

WEISS: That's the yellow of the egg. Most people -- you fell into that trap. Many people try to blend the test questions together, and one of the things they've found --

MERROW: Did I try to do that or did that just happen?

WEISS: Well, many --

MERROW: From racing along?

WEISS: No. Some of the people -- you have many questions to answer. In fact, one of the problems with tests today is that they're trying to answer so many -- they have to distribute all the children in the class along this bell shaped curve, so they have to ask many, many questions, and they only want to give the test for 35 or 40 minutes, so the students can only answer the questions very, very quickly.

MERROW: Speed does count.

WEISS: Speed does count. Well, when you're taking a test if there's a question which troubles you at all you should just skip

it, because there may be other questions further on which you might not get to. There was a study for the Graduate Records Exam, which is required for graduate school today, in the math section. If they gave 15 -- 25 percent of the students today don't finish the test, so this is very much a speed test. If they gave 15 more minutes on the test, almost everybody would finish it and that would change the scores around, or maybe too many people would get the answers correct and then it wouldn't fit their curve.

REINHARDT: John Weiss, of the Committee for Fair and Open Testing.

MERROW: Ironically, the students you might expect to get the highest scores, the brightest, most creative thinkers, are often among those who founder on tests.

REINHARDT: Dr. Banesh Hoffman, who 20 years ago wrote, "The Tyranny of Testing."

BANESH HOFFMAN SAYS THE TEST DISCRIMINATES
AGAINST THE MOST BRIGHT AND CREATIVE STUDENTS

HOFFMAN: Those tests in that format penalize the deep student, the subtle student, the creative student. And that is as true now as it was when I began saying it.

MERROW: Tell me how. How is a creative, deep, subtle student penalized by taking a machine scored, multiple choice test?

HOFFMAN: He's penalized because you know -- nobody asks what are his reasons for his choices. He may have all sort of subtle reasons that the test maker never thought of. And I can tell you a true story. The question said, "In the burning of gasoline in an automobile cylinder, which of the following does not occur," and the answer they wanted was that what didn't occur was the conversion of matter into energy. Now, I pointed out that if you knew Einstein's formula $E=MC^2$, then you had to realize from that that all the energy that was released came from the conversion of matter into energy, and so certainly that wasn't what was not involved.

I tried the question on a Nobel prize winning physicist, and of course he picked not the answer that was in the key. Now, I challenged ETS to defend that particular question, and their defense -- I'll omit the first part, but the second part of the defense, the crucial part, was this, that a student who was sophisticated enough to know Einstein's formula, $E=MC^2$, was expected to be sophisticated enough to realize he shouldn't use that in answering this question.

MERROW: You mean to say ETS said, well, really sophisticated

test takers know how ETS thinks, and therefore they'd know that we're not that deep?

HOFFMAN: Well, this is the way they were arguing, and of course if you're going to argue, well, the test taker, if he knows too much, is expected to know how much he shouldn't use. I mean you make this into a pretty wild sort of game, and certainly it doesn't take on the aura of science that the testers pretend to have.

MERROW: Well, that after all, would be a valuable skill, to understand how much of your knowledge you're supposed to use. Now, we're talking together and you know a tremendous amount of testing, and you're sensing how much or how little I know, and you're not going to baffle me with all kinds of phrases. That's a skill you're using.

HOFFMAN: Yes, that's all right, but you know, who makes a question on a multiple choice test, especially when it's made by a very fancy organization like ETS? It's done not by one person, but by a committee. One person starts it and it's looked at by someone else, and then they have another session. They may even try the question out and then how is the poor person who is too deep going to know whether that question came from a good committee or a bad committee or what?

MERROW: Are you arguing that there are different kinds of brightness and that ETS exams, that standardized tests in general will identify a certain kind, but the truly bright or the slightly out of sync, swimming against the stream, those bright kids are not identified by standardized testing?

HOFFMAN: I'm saying something like that. I'm saying that multiple choice tests, which have to be done, oh, in a question of a minute or much faster, that those tests can select people who are superficially brilliant and quick witted and don't worry about subtleties. But by the same token, those tests will penalize the students who understand things in depth and see things that you have to think about before you can think of what is the best answer and so on.

MERROW: What ETS says, though, is that there are some good answers and that there will be one best answer. You don't believe that?

HOFFMAN: No. I believe that ETS says that, but there is -- they usually say the answer that they consider best, and there are various alternatives. Now, the alternatives are called distractors, and they are deliberately made to try to entice a student away from the wanted answer.

Now, if you make the distractors not distracting -- for example, if you say which one of these is a color, and you say red -- I don't know -- blotting paper or something, and elephants, and that, well, the obvious thing that's a color is red. The distractors haven't been sufficiently enticing, you see.

Now, when you want to make a question more difficult, then you make the distractors more enticing. The trouble is that when you're making distractors more interesting, enticing, then you, yourself, know which is the wanted answer, and therefore the distractor doesn't seem to be more compelling than the wanted answer. But if now you take what you've just done and give it to someone who doesn't know which is the wanted answer, then he will be puzzled by the distractor and probably pick an unwanted answer.

This happens. I have lectured in various places about multiple choice tests, to groups of teachers, PTA's and so on, and very often the teachers come up to me afterwards with tears in their eyes and say that with some of their children the bright children who really like to be sort of intellectually quarrelsome and look at things deeply, when those children have to take one of these tests, the teachers say, "Now, listen, Johnny, please, do me a favor, don't try to be too clever. Just don't think. Just give them the answers they want. I know you're bright, Johnny." And so on, you see. This is a terrible thing that is being done to education, because Johnny is either told be superficial and cynical and so on, which is something we don't necessarily want to train them to be, or else he has to go against the wishes of the teacher he loves dearly, and he is upset that a teacher could take such a position. He's also upset that the whole educational system takes that sort of position.

REINHARDT: Dr. Banesh Hoffman, retired Professor from Queens College, and author of, "The Tyranny of Testing."

A complete printed transcript of all six parts of this series is available for $6.00. The set of cassette recordings costs $18.00.

MERROW: Write to us at National Public Radio - Education, Washington, D. C. 20036.

REINHARDT: $6.00 for the printed transcript of all six programs. $18.00 for the set of cassette recordings.

MERROW: And our address once more, National Public Radio - Education, Washington, D. C. 20036.

REINHARDT: Options in Education is a co-production of National Public Radio and the Institute for Educational Leadership of the

George Washington University.

MERROW: Support for the series is provided by the National Institute of Education and the Corporation for Public Broadcasting.

REINHARDT: John Merrow is Producer for Options in Education. Mary Rowse is our Associate Producer and Rebecca Goldfield is our Assistant Producer.

MERROW: Barbara Reinhardt serves as co-host and reporter, and we've had engineering assistance from Alan Korbuth, Dennis Herndon and George Achavise, of WDBO, Newark, New Jersey. I'm John Merrow.

REINHARDT: I'm Barbara Reinhardt and this is NPR, National Public Radio.

TABLE OF CONTENTS

	PAGE
INTRODUCTION	1
A STUDENT'S REMEMBRANCE OF THE SAT	1
TEST SOUNDS	1
STUDENTS DESCRIBE THE TEST	2
PROCTOR GIVES DIRECTIONS MEANT TO CIRCUMVENT CHEATING	2
A STUDENT DESCRIBES METHODS OF CHEATING	2-3
PROCTOR EXPLAINS HOW TO AVOID GETTING LOST ON THE TEST	3
A STUDENT WHO GOT LOST	3
A STUDENT TELLS OF AMBIGUITY IN THE TEST QUESTIONS	4
ERNEST KIMMEL OF ETS TEST DEVELOPMENT DEPARTMENT GOES THROUGH A SAMPLE QUESTION WITH JOHN MERROW	4-6
ONE STUDENT WHO DID WELL ON THE EXAM HAD HIS SCORES ANNOUNCED IN CHURCH	7-8
ANOTHER STUDENT DID POORLY, AND FEELS THAT IS THE SOLE REASON SHE DIDN'T GET INTO THE SCHOOL OF HER CHOICE	8-9
WARNER SLACK SAYS THAT IN ALL FAIRNESS THE SCHOLASTIC APTITUDE TEST OUGHT TO BE CALLED THE SCHOLASTIC ACHIEVEMENT TEST	9-10
REX JACKSON OF ETS	11
A STORY ABOUT A STUDENT WHO WAS CRUSHED BY HIS POOR SCORES	11

"Copyright © 1980 National Public Radio and the Institute for Educational Leadership. This material has been furnished for non profit, non broadcast purposes and maynot be duplicated or otherwise used in violation of the Copyright Law."

STANDARDIZED TESTING, PART THREE

A STUDENT'S REMEMBRANCE OF THE SAT

BODNER: Do you remember the taking of the test? Do you remember the experience, the morning?

STUDENT: Sure. I was scared to death. I was scared to death, absolutely scared to death, and I sat there and I stared at the exams for awhile and just kind of stared at them for a little while and you know, and that was about the end of it. You know, all I -- by the time I took -- after I was about half way through all I wanted to do was get it done. I stopped looking at the questions. I stopped reading the essays. I stopped doing them. I mean I just started filling in little squares.

BODNER: Filling in any square you could find?

STUDENT: Yeah, I made designs and things like that. Honest to God, I made little designs and stuff like that.

REINHARDT: Taking the Scholastic Aptitude Test, the SAT, can be a harrowing experience for college bound high school students. But a lot is riding on that three hour test.

MERROW: Some colleges won't admit you unless your SAT scores are high. No matter how good your grades are.

REINHARDT: Do we place too much emphasis on this one standardized test? Does it really tell us what we think it does? I'm Barbara Reinhardt.

MERROW: I'm John Merrow, for NPR's Options in Education, with the third part of our series on standardized testing. In this report, a close look at the Scholastic Aptitude Test.

REINHARDT: More than a million high school seniors and juniors take the SAT every year. They spend a tension filled morning sitting every other chair, under the watchful eye of a test proctor.

TEST SOUNDS

PROCTOR: The test will be administered this morning. Check the color of your admission ticket. Blue for the SAT, green for the Achievement. You should all have a blue admission ticket. Make sure you're in the right place. I don't think there's any questions on that.

STUDENT: The tension is very thick. From the minute you walk in and everybody's lining up their pencils on the desk, to make sure

they've got four or five sharpened pencils, and they all have erasers and everybody seems very charged, and just nervous about how well they're going to do. Because they -- everybody realizes that a lot depends on how well they do, and that some people react differently. Some people get very nervous, and I think it doesn't help them at all, and then other people get really keyed up and have a lot of energy for the tests.

STUDENTS DESCRIBE THE TEST

BODNER: Tell me about the morning that you took the test.

STUDENT: Well, I can remember that it was cold weather and I had a very bad cold and a runny nose and I didn't have a tissue with me, and it was embarassing. They wouldn't let me get up and go to the bathroom, and I suffered through the whole test. There were a couple of people who asked to be excused and they wouldn't let us go, except during the break, and it got to be very embarassing and very uncomfortable, because I didn't have a tissue to wipe my nose, so what could I do? It was terrible. It was awful.

MERROW: Students who lose their concentration or do badly for some other reason may repeat the SAT. In fact, they may take it as often as it's offered, usually a half a dozen times a year.

REINHARDT: It costs $9.75 each time.

MERROW: The procedure doesn't vary. That's part of what makes it a standardized test.

A PROCTOR GIVES DIRECTIONS MEANT TO CIRCUMVENT CHEATING

PROCTOR: Place your admission tickets on your desk. If you have any papers of any kind, books, pamphlets, slide rules, compasses, rulers, dictionaries, calculators, please remove them from your desk now and place them on the floor. You may pick them up after the exam. Close all your purses. Place them on the floor next to your desk. Please do not open them until the conclusion of test. No student may leave until I announce the dismissal.

A STUDENT DESCRIBES METHODS OF CHEATING

STUDENT: There's a tremendous amount of cheating on the SAT's. It's a three hour exam, and after each hour there's a five minute break. Students go out of the room, go to the bathroom. Many students bring dictionaries and keep them in the bathroom. Many students will sit in the back of the room. You are allotted half an hour for each section of the exam. There is a test of standard written English and an experimental section, which take up one hour. The SAT per se takes up two hours, and there are two math

sections and two English sections. Many students, even though they're not permitted to do this, will sit in the back of the room, out of view of the proctor, and spend the entire time working on the English part of the exam, all three hours. And then they'll come back and take the test and they'll cancel the math, which they're permitted to do. You can cancel either section.

Then they'll come back another time and spend three hours working on the math part of the exam and cancel the English. And of course the dictionary is in the bathroom, being able to discuss with students the scores. I think that it's very inequitable and there are so many variables, so many ways of cheating that it's very unfair, I think, to weigh the SAT's too heavily.

REINHARDT: Harvard freshman, Kenny Drexler. Students are allowed to cancel one portion of their test, but less than one half of one percent actually do.

MERROW: Last year, Educational Testing Service, which makes up and administers the test, investigated 1,400 cases of possible cheating. The most common forms? Copying or having someone else take the test for you.

REINHARDT: ETS says it's very hard to cheat. That's one of the best things about the test, they say.

A PROCTOR EXPLAINS HOW TO AVOID GETTING LOST ON THE TEST

PROCTOR: On your answer sheet there are 50 numbered spaces for each section. Mark your answers by darkening the appropriate spaces on your answer sheet. Make sure that all of your marks are dark and that they completely fill the spaces. Do not make any stray marks on your answer sheet. If you erase, do so completely, because incomplete erasures might be read as intended answers. Be sure that the number of every space you darken on your answer sheet directly corresponds to the number of the question you are answering. You may work out your answers in any --

A STUDENT WHO GOT LOST

STUDENT: I left a question blank and then as I went on I kept filling in, and I realized I think it was like five questions later, this was when I was taking the SAT, that I was on the wrong question. I had to go back and retrace my steps, because those bubble sheets are so hard to figure out. And it's confusing for your eye, to know exactly which line you're on.

MERROW: He was lucky that he caught his error. If he hadn't, he might have gotten every question wrong.

REINHARDT: The SAT is divided into a math and a verbal section. The maximum score for each is 800, for a maximum composite score of 1600.

HERROW: There are 60 math questions and 85 verbal questions, all of which are multiple choice.

STUDENT TELLS OF AMBIGUITY IN THE TEST QUESTIONS

BODNER: Have you ever had this experience, where you're looking at a question and you can see a perfect justification for two or three of the answers, and you just want to say -- you want to scream out and say to somebody something like, "Listen, just let me explain and I'll tell you why?"

STUDENT: Oh, sure. I mean that happens all the time. You know, especially like if it's the GMAT's or the SAT's, when you're reading an essay and you're evaluating the essay, and you look at this and say, well, that could be right and that can be right and, you know, that can be right too. And they're assuming that there's a black and white, and one of the major lessons you learn in life is there's nothing that's black and white. There's always a gray area. And --

BODNER: So what do you do with that frustration?

STUDENT: Well, you just kind of sit there and go, well, it's -- then you guess. Then you're making what they call an intelligent guess. You know, intelligent guess. You sit there and your strategy is to go through the questions and say, well, which one is the least likely, and then you keep -- you cross it off until you find which is the most likely, which is absurd, because everybody's thinking processes are different.

REINHARDT: Test makers write the questions so that most, maybe all, of the answers are plausible. The nearly right answers are called distractors. Ernest Kimmel, head of Test Development for Educational Testing Service, walks John through one of the questions.

ERNEST KIMMEL OF ETS TEST DEVELOPMENT DEPARTMENT
GOES THROUGH A SAMPLE QUESTION WITH JOHN HERROW

KIMMEL: Aesthete is to beauty as an enthusiast is to a cause, hunter is to nature, administrator is to government, advocate is to legality, or philanthropist is to money. We have a noun that might be used to describe a person in relationship to a particular attribute that that person would be interested in. An aesthete, someone who is interested in beautiful things, aesthetic things. And the task of the student facing this question is to find someone who has that same kind of relationship.

MERROW: There's no easy throwaway among these five choices.

KIMMEL: This is a difficult one which is up there to help sort among the students who are very good and the ones who are really super. So, you know, let's go through them. An enthusiast is -- sort of adopts a cause, does something to promote a cause. The hunter is out in nature, but some people might perceive him as acting upon nature, doing something to affect nature. I'm never quite sure what administrators and government do with each other, but they do seem to go together. The administrator is, in fact, one who works in government, who carries out the policies of government. An advocate clearly deals with the legal system. Legality, does he admire legality in the way that an aesthete admires beauty?

MERROW: Does an advocate exist in relationship to legality the same way an aesthete exists in relationship to beauty.

KIMMEL: That's the question the student has to ask, or the philanthropist who gives away his money, or does something in the disposing of money. Now, the aesthete does not dispose of beauty, so I think we can perhaps do that, and probably the aesthete does not take part in beauty. He himself is not part of beauty, as the administrator might be of government. So we can eliminate that. The hunter, again, I think we can probably eliminate that. The hunter acts upon nature in a way probably quite different from the relationship an aesthete feels to beauty. My own thinking, I'd come down and I'd ponder between the enthusiast to a cause and advocate to legality, thinking of legality as some great concept of justice, you know, lady justice holding her scales, or whatever. And try to decide between those two things.

As we're talking, I would choose alternative A. What would you choose?

MERROW: I would choose B. I would choose the hunter's relationship to nature. Or at least I would want to make a case that one could choose that, depending on your attitude toward hunting. If you see hunting as kind of a cruel and destructive, which I guess is somehow your feeling, and probably is mine also, but many people see hunting as an essential part of man's relationship to the world as maintaining some kind of balance, as a necessary kind of thing, and as a productive and fulfilling act. I mean couldn't one argue that B could be the right answer, depending on where you were coming from in relationship to hunting?

KIMMEL: Okay, but even accepting that hunting has its socially useful purposes, the hunter is somehow altering nature in doing that, and I don't think you can argue that the aesthete is altering beauty, which is some ethereal concept that he's respecting

or worshiping.

MERROW: Well, I think you could. You could say yes, this is -- the aesthete also defines beauty, and if in fact is an artist then shapes and determines beauty.

KIMMEL: Well, I think the aesthete is one who recognizes and admires beauty, not that he's doing anything with it. Other people may have aesthetic tastes, but the aesthete, by its meaning, is one who recognizes and admires beauty, not that he necessarily creates it or shapes it.

MERROW: I'm not trying to nitpick. I really would have a lot of trouble choosing among enthusiast's relationship to cause, hunter to nature and advocate, legality. You would have trouble choosing between A and D. Do you know the right answer?

KIMMEL: I would choose A, the enthusiast sort of recognizing, admires a cause, as the best relationship here. Shall we look and see if that's --

MERROW: By all means.

KIMMEL: If that's how it was scored.

MERROW: We're lucky the answers are in the back. You got it right.

KIMMEL: I got it right. Only about 20 percent of the typical student taking the SAT gets that one correct.

MERROW: But that's a good question, though, if 20 percent --

KIMMEL: It's a good question. It's a difficult question, and one which only students who are quite able to deal with not just the first concept, because immediately all of the alternatives bears a relationship within the analogy. But that has to be a student who can push beyond that to finer meanings or finer relationships within the context of the analogy.

REINHARDT: Ernest Kimmel, who helps develop the SAT for Educational Testing Service. By the way, on the average you have less than a minute per question to come up with your answer.

MERROW: The Scholastic Aptitude Test was created 60 years ago, by the College Entrance Examination Board. The goal was a standard exam that all colleges could rely on, instead of each having to give its own entrance test.

REINHARDT: In 1947, the College Board and a couple of

philanthropic foundations created the Educational Testing Service, then hired it to run the SAT.

MERROW: Today, most colleges do use the SAT score. Some use it as a cutoff and don't take anyone below whatever point they've chosen, ignoring the fact that the SAT score is not that precise.

REINHARDT: ETS recommends that colleges look at grade point average, rank in class, letters of recommendation, the student's own essay and the SAT score.

MERROW: In fact, grade point average is a better predictor of how a student will do in college.

REINHARDT: Few people know that all the SAT is supposed to do is to predict a student's performance in his first semester. SAT scores have taken on a greater importance than was intended.

ONE STUDENT WHO DID WELL ON THE EXAM
HAD HIS SCORES ANNOUNCED IN CHURCH

STUDENT: Gosh, I remember going to church and having them make a formal announcement about it in church, that Dave Fleming had gotten 1360 on his SAT's.

BODNER: Wait, wait. In church? What does church have to do with SAT's?

STUDENT: I guess it was the closest thing to initiation or steps of manhood that a lot of the church members, who are lawyers and stuff like that, they felt like it was my first step into manhood.

BODNER: So you got this mid-1300 score and they announced it after mass?

STUDENT: Everyone thought I was a gifted child after that. Then they felt like, oh, my future is secured. I can now introduce you to my little daughter, you know, who is 15 or 16. You know, look him over. You know, in a few years he'll be successful. And so it was almost as if I was now a breeding species for some of the people in church.

BODNER: The SAT as a rite of passage?

STUDENT: Yeah. Very much. Because often there wasn't very much else that you could put actually figures on. I mean everyone else might get good grades, but it's a matter of how do you relate nationally, okay? And many of the guys who are lawyers or doctors or something of that nature were looking on kids who had some sort

of basic strong desire to do well. They really felt like the SAT's were the first sign of good hope.

I found that life was a lot easier. It was kind of like getting your key into the country club or something like that. At least in that circle of people. I don't even believe they remembered my name before that time.

BODNER: How did the church ever find out about your SAT's? I thought that's supposed to be a strictly academic matter.

STUDENT: Mother announced it. Mother announced it because it's a pretty important thing around that church.

BODNER: Suppose you hadn't done well?

STUDENT: It wouldn't have been mentioned. I would have tried to go through college. I might have gone out with their daughters anyhow, but they certainly wouldn't have been looking for me. And for them SAT's were the first sign of wealth or possible potential wealth.

REINHARDT: Low SAT scores have a life of their own as well.

ANOTHER STUDENT DID POORLY AND FEELS THAT IS THE SOLE REASON SHE DIDN'T GET INTO THE SCHOOL OF HER CHOICE

STUDENT: I had an interview at one school, a wonderful interview that lasted about a half an hour, and we talked about all aspects of my life, what my interests were, and then the subject of SAT's came up and she said, "How did you do?" And I said, "Low 500's." And then I said, "What are my chances of getting in here?" And she said, "Well, it's a borderline because of your SAT scores." and I was surprised, because I'd had such a good interview, and I thought that I'd made a good impression and that she was impressed with me.

REINHARDT: Do you find that teachers, professors or, now that you're considering transferring to other schools, do people still ask you what you got on your SAT's?

STUDENT: I've noticed in certain areas those questions are asked. I know that when I was interested in trying out for the swim team, I received a questionnaire and one of the questions was what were your SAT scores. And it does keep cropping up. And the other side of it is, too, that I'm a student in good standing at the university, and so the discrepancy still holds. I have those awful SAT scores and I have good grades to prove that I can do otherwise. I have qualities that are worthwhile and the SAT's don't tell you that. But I think that that is something that's very important to

consider in accepting people to college and to graduate school.

I know that I have a friend who is doing extraordinarily well where she is, and she's upset because she's not studying as hard as she should be but doing well. And people are constantly hounding her and saying "Well, you don't have to worry. You don't have to kill yourself because you did well on your SAT's." So it really does follow you. It's kind of a stigma. You get a 600 and you're in one class. You know, you get a 300 and you're in another class. And it seems that until you prove otherwise, if there is a way to prove otherwise, that it does follow you around.

REINHARDT: One reason the Scholastic Aptitude Test, the SAT, is under attack today, is that word aptitude, which implies innate capacity to learn, intelligence. Dr. Warner Slack.

WARNER SLACK SAYS THAT IN ALL FAIRNESS THE SCHOLASTIC APTITUDE TEST OUGHT TO BE CALLED THE SCHOLASTIC ACHIEVEMENT TEST

SLACK: I'm against tests that purport to meaasure aptitude or intellectual calibre or the capacity to learn, when they don't. I'm really more concerned about the self-esteem of the student. I think it's more important how the student feels about her own mind than where she goes to college.

If a student with a low SAT score believes that this score accurately reflects his intellectual calibre or his capacity to learn or his aptitude and that no special preparation or training could have altered his performance, then his self-esteem has been endangered and his mental capacity has been disparaged, and if the message that special preparation and training doesn't work is false, as it clearly is, then that student has been seriously wronged, and in my view students in the United States have been seriously wronged by the Educational Testing Service.

MERROW: I remember when I was a high school senior, and I graduated in 1959, a kid who was a tremendous student, a capable guy, good athlete also, did poorly on the SAT's when we all took them, and there was a kind of a taint. Somehow he was different. You know what he was labeled? He was labeled an over-achiever. Remember that?

SLACK: Right, yeah. As a matter of fact, nowhere really is the undeserved clout of the SAT more evident than in the use of this word over-achiever, because here you have students who get good grades, who have performed well in the area that predicts best, that correlates best with college performance, and then have a bad day with a test and they're called over-achievers, if they somehow had transcended their intellectual potential. And all we know about the

SAT is how well it correlates with first year college performance. That's all that's known, and it's a very poor predictor, not nearly as good as high school records.

MERROW: As if the four hours was the true picture and the four years was somehow false.

SLACK: Exactly. Exactly. It's terribly unfair.

MERROW: Stories about kids whose attitudes have changed about themselves because of their SAT scores. The other side could be true. Kids who do well might have a distorted picture of their own value and somehow feel they're much better than the rest of the world.

SLACK: That's a very good point. I think that high scores may well be hurtful to students if they haven't been accomplishing in other areas. They may think they don't have to buckle down. There's this strange situation in the United States where doing well without having prepared or studied is somehow helpful to one's prestige. It's like in college students pretending they haven't studied for a test and then going in and doing well, to make it appear that they're smarter. And the competition for SAT scores is, I think, of the same ilk. We all use tests every day, all the time. I just want the test to be fair and open. I want the student to recognize the fact that he can study for these tests. I want him to see the correct answers after he's taken the test. And I guess the best test would be the one where the study for it would prepare one not only to do well on the test, but to move on and do well in the future effort, like in college.

MERROW: You're never going to make a test maker, then, because a good test item is one that not everyone gets right, that only some of the class gets right because that allows you to discriminate among the folks in the class.

SLACK: Well, that's one view. Mine is different. I like the test items that everybody can get right. I think our culture is too hung up on prophesy. That's a legacy of Plato, I think, where prophesy is more important than accomplishment. I much prefer the situation where we want every student to get 100 if they possibly can.

REINHARDT: Dr. Warner Slack. He and Dr. Douglas Porter, both of Harvard Medical School, are the authors of, "The Scholastic Aptitude Test, A Critical Appraisal."

MERROW: Even though the word aptitude is in the test's name, ETS now is careful to call the SAT a measure of general achievement. Rex Jackson of ETS.

REX JACKSON OF ETS

JACKSON: What we have here is a measure of academic ability that is not an infallible kind of indicator. It's not a prediction or some kind of ceiling on how you're going to perform or succeed later on, but it's simply one aspect of the description of you as an individual, along with the other information that is an important part of the college admissions process.

MERROW: In effect, you're kind of washing your hands, saying look, don't blame us if anybody over-relies on it.

JACKSON: No. I think it's very important for us to continue to educate and communicate to individuals what the scores mean and how they can properly be used. But I'd sometimes like to turn that question around, because I don't know what the alternatives are. Should we discard any kind of sound measure of academic achievement because it might be misused? Should we do away with the SAT, which is in fact one of the most reliable in a technical sense measures that's used in the college admissions process, because it might be misinterpreted? If some students could on occasion beel bad because they didn't score as well as they would, does that mean that we shouldn't have an academic standard that the students should aim for? Should there be no kind of evaluation? It's -- there are -- there are certainly unhappy aspects of this kind of process, and sometimes students will not be satisfied with what they achieve. But I don't think that's an argument for going to a system where we try to make everybody happy.

REINHARDT: Rex Jackson of Educational Testing Service.

STORY ABOUT A STUDENT WHO WAS CRUSHED BY HIS POOR SCORES

STUDENT: There was one guy who took about three or four -- I remember about three or four study courses for the SAT's. They were -- I don't know who they were run by, but his father really wanted him to be a success all the way, all over. He choked. He really choked. It was out of his comfort zone. By the time he got there he just froze, and the parents didn't make much to-do about it. They sent him to UVA, which is a nice, average, typical school for Northern Arlington kids to go to. They -- he wasn't discussed much after that in the household. You know, he was a good friend of mine, but no longer were we talking about law school. We were talking about -- you know -- he was talking about getting through four years, working with his father, you know, and hopefully, you know, if his father passed on he might get the business, but certainly not something where he would be a genuine success the weay he wanted to. And I feel that was really a doggone shame, because he's a brilliant kid. It really shattered him.

REINHARDT: A complete printed transcript of all six parts of this series is available for $6.00. The set of cassette recordings costs $18.00.

MERROW: Write to us at National Public Radio - Education, Washington, D. C. 20036.

REINHARDT: $6.00 for the printed transcript of all six programs. $18.00 for the set of cassette recordings.

MERROW: And our address once more, National Public Radio - Education, Washington, D. C. 20036.

REINHARDT: Options in Education is a co-production of Natonal Public Radio and the Institute for Educational Leadership of the George Washington University.

MERROW: Support for the series comes from the National Institute of Education and the Corporation for Public Broadcasting.

REINHARDT: John Merrow is the Producer of Options in Education. Mary Rowse is our Associate Producer, and Rebecca Goldfield is our Assistant Producer.

MERROW: Barbara Reinhardt serves as co-host and reporter, and we've had engineering assistance from Greg Hedgepath and Terry Knight.

REINHARDT: Some of the material for this program came from Pat Bodner in Boston and Greg McVicar in Sacramento, California. I'm Barbara Reinhardt.

MERROW: I'm John Merrow, and this is NPR, National Public Radio.

Program No. 283: Standardized Testing, Part Four

TABLE OF CONTENTS

	PAGE
INTRODUCTION	1
MATH SESSION OF A COACHING COURSE	1
STANLEY KAPLAN, WHO OWNS 90 COACHING CENTERS	2
KIDS WHO ARE TAKING THE COURSE	2-3
ARTHUR LEVINE, AN ATTORNEY WITH THE FTC WHEN IT INVESTIGATGED THE MERITS OF COACHING	3-4
ONE STUDENT DIDN'T COACH FOR THE SAT BECAUSE SHE HAD BEEN TOLD SUCH STUDY WAS USELESS	4
ACCORDNG TO WAQRNER SLACK, ETS OMITTED IMPORTANT DATA WHICH SHOWED THAT COACHING WAS EFFECTIVE	5-6
REX JACKSON OF ETS SAYS ETS NEVER DENIED THE EFFECTIVENESS OF PROPER COACHING TECHNIQUES	6-8
JOHN WEISS EXPLAINS THE CONCEPTS OF 'RELIABILITY' AND 'VALIDITY' IN TESTING	8-10
WILLIAM BOYD TELLS OF LOW-SCORING SAT TAKERS WHO WENT ON TO TREMENDOUS SUCCESS	10-11
AN ADMISSIONS DIRECTOR WHO STILL FAVORS STANDARDIZED TESTING	11-12
WILLIAM TURNBULL, PRESIDENT OF ETS	12

"Copyright © 1980 National Public Radio and the Institute for Educational Leadership. This material has been furnished for non profit, non broadcast purposes and maynot be duplicated or otherwise used in violation of the Copyright Law."

STANDARDIZED TESTING, PART FOUR

MATH SESSION OF A COACHING COURSE

TEACHER: X has positive values. Y has negative values. If I add X and Y together, what happens? Okay, it's going to depend upon what the choices are for X and Y. We want the smallest possible value for that --

MERROW: She's coaching high school students for the Scholastic Aptitude Test, the SAT.

REINHARDT: The room is packed with anxious students, determined to raise their scores. They'll meet every Thursday night for 11 weeks and pay $300.00 for the course.

MERROW: That $300.00 may well raise their scores and ease the passage to the colleges of their choice.

TEACHER: X can go from 3 to 7, including 3 and 7, and 3, 4, 5, 6, 7. Look at your choices for Y. What are the values that Y can have? Negative 3, negative 4 and 5.

REINHARDT: What about students who can't afford $300.00? Is it fair for colleges to consider scores money can buy?

MERROW: There are other questions of fairness too. In the SAT itself and in the way it's used. I'm John Merrow.

REINHARDT: I'm Barbara Reinhardt, for NPR's Options in Education, with the fourth part of our report on standardized testing.

TEACHER: On the SAT you'll have 50 questions, and the questions will be of two types. Type one will be on usage. Type two will be on sentence format. Relax, there won't be any writing or composition type of things required on the standard written English test. Instead, what you'll have to do is recognize mistakes in grammar. Your scores will range on this one --

REINHARDT: Coaching kids to take the SAT is big business. You can pay $7.95 for a book, or up to $400.00 for a course. Even the test maker itself, Educational Testing Service, ETS, provides students a copy of an old exam to study.

MERROW: The biggest coaching operation of all is owned and run by Stanley Kaplan. His 90 Kaplan Educational Centers coach 50,000 kids a year and bring in $15 million.

REINHARDT: Kaplan doesn't call it coaching. He prefers the term

supplementary education.

STANLEY KAPLAN, WHO OWNS 90 COACHING CENTERS

KAPLAN: Actually, when people come to us, they come to us to improve their scores. But whether they know it or not, they become an improved student, and because they're becoming an improved student, that's why they're getting a better score. So therefore, my range is to bring back, resurrect their background, to hone their skills, to get them to learn to think innovatively, instead of just memorizing and regurgitating.

HERROW: When someone signs up here, to be coached or to be re-educated, whatever you want to call it, do you guarantee their scores will get better?

KAPLAN: Absolutely not. In fact, you know, we have pretty good results, but when students call us up, you know, prospective students, and ask us what kind of improvement, what our track record is, we don't give it, because, you see, say our average, our average improvement on the SAT is about 100 points, math and the sum total, math and verbal. But if we tell the student our point improvement is 100 points, one of our students may go up 200. Somebody else will go up zero. There's your average of 100. But the person who went up zero, he's not worried, he's not concerned about the guy who went up 200. Where's that 100 points you promised me?

So statistics works for an average. It doesn't work for any individual.

HERROW: 50,000 students. $300.00 a pop. Is Stanley Kaplan Educational Centers, Incorporated a $15 million a year business?

KAPLAN: $15 million a year, as far as total receipts, yes.

HERROW: What does that tell you about this society's interest in test scores?

KAPLAN: Well, I get another picture. Society is interested in test scores. Society is interested in learning.

HERROW: Which is it?

KAPLAN: Well, I would say half and half

KIDS WHO ARE TAKING THE COURSE

STUDENT: I don't think the SAT's are a good measure of much of anything. I really think they're a waste of time and silly.

BODNER: So what does this test actually measure about you? Does anyone have any ideas?

STUDENT: I think the way you can cope with your surroundings and ignore everything that's going on, because even if you're sitting there and you're hungry, your stomach is saying feed me, and you're trying to concentrate on this marathon test, and you just can't do it, because even the littlest things turn you into -- enter into your subconscious, and it just messes you all up. You can't concentrate.

STUDENT: It's not that important. It's just you worry about whether you can really get into a really good school. The competition is just so stiff that you are worried all the time. Like your senior year, you're a nervous wreck. At least I know I am. So it's just -- it's nerve wracking, is what it is.

STUDENT: A lot of it's mental, if you believe yo'll do better by taking these classes. Then you may do better and your mental aspect will be better.

STUDENT: I've already taken it twice, but I'll be even more at ease now, because not only do they prepare you academically and really physically for the task, but they tell you what to do. Like they say take a candy bar along and you'll get energy and you won't be distracted by the fact that you're hungry, or make sure that you're comfortable in temperature and all sorts of things like that.

MERROW: Students at a Stanley Kaplan Educational Center. They're each paying $300.00 for a coaching course, hoping to improve their scores on the Scholastic Aptitude Test.

REINHARDT: Coaching can work. Even the test maker, Educational Testing Service, admits that.

MERROW: It took years to settle the question, though. With ETS on one side, the coaching schools on the other.

REINHARDT: In the late '70s, the Federal Trade Commission decided to investigate. Arthur Levine was then an attorney with the FTC.

ARTHUR LEVINE, AN ATTORNEY WITH THE FTC
WHEN IT INVESTIGATED THE MERITS OF COACHING

LEVINE: Well, there seemed to be a real paradox. On the one hand you had the Educational Testing Service saying that coaching was ineffective, it did not good. But yet on the other hand you had the coaching schools representing that they did do a significant job. And too, you had a lot of people going to coaching schools.

So just that bit of evidence, the fact that people were flocking to them tended to indicate that they would have a positive effect. Why else would anybody spend 250 or $300.00? They're certainly not going to give it away, so we found -- we decided that someone had to be wrong. We didn't know who it was when we initially went into the venture.

REINHARDT: What were your suspicions?

LEVINE: My suspicions were that the coaching schools were wrong. I thought that ETS's literature was accurate. And I probably thought that because I had always been a beneficiary of standardized testing. I had done reasonably well and gotten into the college I wanted to go to, went to the law school I wanted to go to. At least in part as a result of scores on standardized examinations. So I guess I had a vested interest in hoping that they did really measure something because I had always done fairly well on them.

MERROW: In his draft report, attorney Levine concluded that coaching could raise SAT scores significantly, as much as 100 points out of a maximum of 800.

REINHARDT: But the FTC decided not to publish Levine's report. Instead, it reanalyzed the same data and came up with a different conclusion, a coaching program that lasted at least 40 hours could lead to a 20 to 35 point increase.

MERROW: A year later, in May, 1980, the Educational Testing Service came publicly to the same conclusion, coaching can raise scores.

REINHARDT: The official pronouncement must have shocked a lot of people who had believed all along that the Scholastic Aptitude Test was a true picture of what they were, not just another test to study for.

ONE STUDENT DIDN'T COACH FOR THE SAT
BECAUSE SHE HAD BEEN TOLD SUCH STUDY WAS USELESS

STUDENT: I felt that the nature of the name, Scholastic Aptitude Test, meant that I was to take the test and use whatever skills I had acquired in order to pass the test. And I didn't feel that coaching was something that was a valid thing to do. I'm certainly not putting down people who did take the coaching, but I feel that what I learned should have been sufficient to help me to do well on the test.

MERROW: Her combined score on the math and verbal sections o the SAT was around 900, out of a maximum of 1600. She was an A student in high school.

REINHARDT: About the same time that Educational Testing Service was admitting that coaching can improve SAT scores, another report was published, this one accusing ETS of years of systematically and deliberately ignoring research studies that contradicted its official position.

MERROW: Basically, the authors, two Harvard Medical School doctors, accused ETS of fraud. Dr. Warner Slack.

ACCORDING TO WARNER SLACK, ETS OMITTED IMPORTANT DATA WHICH SHOWED THAT COACHING WAS EFFECTIVE

SLACK: There's ample evidence to indicate that training or special preparation for the SAT can be highly effective in helping students to raise their scores. And perhaps most disturbing in regard to the coaching issue is that Educational Testing Service and the college board have omitted and misrepresented the data on coaching. They have ignored studies that showed large gains due to coaching, studies they knew about.

One study, the Merin Study, done by a West Point investigator, was supported by the college board, and it showed sizable gains due to coaching. And ETS and the college board omitted reference to this study repeatedly.

MERROW: But now they don't. Now they're acknowledging that there are ways that students can make themselves better prepared for the tests.

SLACK: They're changing their tone in response to the criticism, but the facts are there. The history is quite clear that they have misled students, their teachers, their counsellors, their parents and college admissions officers, with their claims that coaching doesn't work.

MERROW: So you conclude from that that the kids who went ahead and got coaching anyway somehow had an unfair advantage.

SLACK: Yes. The issue of coaching is particularly important, because students who believed the claims of Educational Testing Service and didn't prepare for the test may have been needlessly deprived of admission to the college of their choice.

But more important, really, to Doug Porter and me is the fact that a student with a low SAT score who believes that this score is an accurate reflection of his aptitude or intellectual caliber as ETS has called it, this student's self-esteem has been weakened and his mental capacity has been disparaged, and this is highly inappropriate.

MERROW: You really do feel, though, that it's a damaging

thing to get back the 350 or the 450 when others around you are scoring much higher?

SLACK: Yeah. It's clear that students compare scores and make judgments about each other's intelligence on he basis of the SAT. It's a common conception that the SAT measures intelligence. Aptitude, when you look up the word in Webster's, is the capacity to learn. Intelligence is defined as the ability to learn. I would defy the verbal item analogy writers at Princeton to tell the difference. And they've used other phrases, like intellectual caliber, and they've told students that you can't study for the test. That implies that they're measuring something that is insensitive to experience or training. And the child who believes this, and the child who gets a low score, the student who gets a low score and believes that no training or special preparation could have altered his performance, he is -- his self-esteem has been endangered. And if that message is untrue, which it clearly is, we have a lot of data to show that, then that student has been seriously wronged, and I think Educational Testing Service has broken the student's trust.

REINHARDT: Warner Slack. He and Douglas Porter are authors of, "The Scholastic Aptitude Test, a Critical Reappraisal."

MERROW: Educational Testing Service says the real issue is the amount of time involved in coaching, whether it's short term cramming or long term instruction. But ETS doesn't say when short term becomes long term.

REINHARDT: We asked Rex Jackson of ETS whether saying coaching can work amounts to an admission ETS was wrong before.

REX JACKSON OF ETS SAYS ETS NEVER DENIED THE EFFECTIVENESS OF PROPER COACHING TECHNIQUES

JACKSON: No, I wouldn't say so, because we've also said that the tests are intended to measure developed abilities and that sound educational experiences contribute to that kind of development, and that's been said clearly. So I think what we're saying is really consistent with what's been said over the years.

MERROW: Well, it's not hard to find ETS statements about coaching, from documents over the years. Here's one from 1972. "If this thesis that scores can be improved were to be proved true, its proof would challenge the validity and essential purpose for which the test was constructed." That's 1972.

Another reference to coaching as, "...this obnoxious practice." As, "...a subversion of the educational process."

Here's another. "The abilities measured by the SAT develop over a student's entire academic life. So coaching, vocabularly drill, memorizing facts or the like can do little or nothing to raise the student's score." There are a lot of those.

JACKSON: Yes, and those are all out of context, also. That first quote, by the way, is, I believe, I recognize as a statement from a research report that was written by Dean Whitlow at Harvard. I have a statement here that goes back to 1957, which I'd be prepared to read parts of if you'd be interested in, because I think it gives a broader picture of what we said about the test. This is from a student booklet on the SAT from 1957.

It says, "Advice on how to prepare for a Scholastic Aptitude Test probably would best be given to students early in life. Do your lessons conscientiously. Read widely and thoughtfully. Observe the world around you. Test what you are learning against the world as you see it. All that is perfectly good advice, because mental abilities, like physical abilities, are products of growth, practice, instruction and experience over a period of time. But you are now planning to take the SAT soon. What, then, can you do now to prepare for the examination? First, a don't. Don't try to review in a night or a week all you've done." And it goes on in that vein. It refers to that kind of preparation as a treasure hunt.

I'd just like to read a final part of the statement. "The treasure hunt type of preparation more commonly is called cramming. Coaching is another word for it, though coaching may be a sound and rewarding method. If coaching means an honest effort under the guidance of a good teacher over an extended period of time, to improve those skills in which you are weak, then it can be recommended as effective study. But coaching courses or coaching schools which promise to give you a collection of nuggets of fact or to provide intensive drill with sample questions that will, quote, assure your success, unquote, are not worth your time or money."

I think that's the kind of distinction that's been made over a long period of time.

MERROW: Does it say anywhere in these guidelines that SAT has a standard error of plus or minus 30 points?

JACKSON: I don't believe it's in the guidelines, because these are general principles. There are interpretive booklets that include information for interpreting scores more specifically.

MERROW: I talked to a lot of kids. I have never met a kid who has said, "Well, I got a 430, but it could be the same as a 460," or had that kind of understanding of a score. Have you?

JACKSON: I don't recall that. I also haven't talked to students who said, "My grade point average in high school was 3.1 but it might have been 3.3 or 2.9 if we took the error of measurement." I think they do have a conception, when you see a number, that it's a rather precise and fixed number. But all of these things that are dealt with in admissions, test scores, grades, certainly information from recommendations and so forth is subject to error.

REINHARDT: Rex Jackson of Educational Testing Service.

MERROW: A standard error of measurement is built into every standardized test, including the SAT. I said plus or minus 30 points, but others say it's 38 points, even 50 points.

REINHARDT: Think of it this way. If I say meet me at 4:30, in my mind I take into account a standard error. I know you might come at 4:25 or 4:35, and I won't be upset. To me that is the same as 4:30.

MERROW: The SAT score works the same way. A score of 500 is really no different from a 525 or a 540.

REINHARDT: There's another piece of the equation, called statistical significance.

MERROW: A test like the SAT is significant to a certain percentile, the 90th or the 99th.

REINHARDT: Let's say you and I are supposed to meet every day on the corner at 4:30. Well, after awhile I'll be able to calculate the probability of your appearing around 4:30. If this test is significant to the 90th percentile, that means that 90 of the time you'll get there on time. The other 10 percent of the time, you'll be too early, too late or you won't show at all.

MERROW: The SAT is significant only to the 90th percentile. that means that 10 percent of the time you'd get a vastly different score from your standard error range. Measurement experts use the terms reliable to describe the standard error of measurement and the level of significance. John Weiss.

JOHN WEISS EXPLAINS THE CONCEPTS OF 'RELIABILITY' AND 'VALIDITY' IN TESTING

WEISS: Reliability refers to how well the test agrees with itself. For instance, let me just give you an example for the SAT. Students' reliability. They have a .9 reliability on the SAT, and that sounds very impressive, and ETS tell you that. But most people don't understand that this means the following: That standard

deviation is 30 points on the SAT. So if you score 550 on the SAT, that means nine times out of ten if you take the test again and there's no learning that's occurred, you take it the next day and you forgot everything you learned when you took the first day, your score will be between 500 and 600. That's how reliable the test is.

Now, a 500 on the SAT is about -- used to be, when they made the test, was the 50th percentile, while 600 was the 84th percentile. That's a big difference. And you take 90 percent of the students, when they retake the test, will fall somewhere in this big range.

MERROW: So if I got a 550 on my SAT's, I really ought to, somehow in my heart, know that, hey, I'm somewhere between 500 and 600.

WEISS: Yes.

MERROW: But anybody looking at it is going to say, hey, Merrow is a 550.

WEISS: Exactly. But then you have to realize again that 10 percent of the students who retake the test, and this is without coaching or any learning occurring, are going to score either -- 5 percent will score above 600, which goes -- will get you into almost any college in the U. S., and 5 percent will score below 500, which means you're a candidate probably for a not very competitive college. And this is by luck. And so in addition one of the things they've found out -- this is according to ETS's own publications, which they don't publicize, mind you, but we dug up, that when you take the test and then you take it again, because you're familiar with the test and the questions, your score improves on the average of 30 points.

MERROW: Validity.

WEISS: Validity is how well the test is supposed to measure what it purports to measure. The SAT, they use first year predictive grades. This is only first year of college grades. Over four years how many students graduate, there's almost no prediction, and how well they're going to do after they graduate is almost zero prediction. A man named -- a Harvard professor, David McClelland did a longitudinal study of people that graduate from many, many different colleges, everything from prestigious Ivy League colleges to state institutions, and followed them throughout their careers. And these are people who scored, you know -- this was their grades and test scores who did very high and very low. 25 years later you couldn't tell them apart.

Now, another aspect of this test is after -- many students who come from disadvantaged backgrounds, the first year of

college is the hardest, especially if they're going to very hard institutions from different socio-economic status. They're in a culture shock. They don't know how to study. They may be in a competitive program. So they're having difficulty trying to learn -- you know, learning many different things. But the second year the predicted validity for the SAT is even worse. But it's important to remember that while the test does not measure how well you're going to do in college very well, it does measure race and socio-economic class almost perfectly.

REINHARDT: John Weiss of the Committee for Fair and Open Testing.

WILLIAM BOYD TELLS OF LOW-SCORING SAT TAKERS
WHO WENT ON TO TREMENDOUS SUCCESS

BOYD: Basing college admissions on freshman year performance is something that I think is limited, a limited perspective, and I would make an analogy to somebody who would argue that we should award the gold medals at the next Olympics in the 100 meters to the people who were ahead at 20 meters. We don't do that. We award it to the one who is ahead at 100 meters, at the end.

REINHARDT: William Boyd uses a lot of his energy trying to persuade colleges to look beyond, maybe even disregard the SAT scores of the kids his program sponsors. He describes what can happen.

BOYD: We have, one, the most dramatic example which I will use just under that label, because it is by no means typical. But this fellow came from a tenant farmer's background in a small town in Georgia. He had scores that were so low they barely reached the graph. It was amazing. He went on through one of our schools. He did -- he struggled, he did all right. He got into an Ivy League college. He was one of the first blacks to win a Rhodes Scholarship. He has gone on since then, finished law school with distinction, has been clerk to a federal district judge and is now practicing securities law in one of the major -- by major I mean white -- downtown law firms in a southern city. That's the most extreme example I could give you of where in any isolated case the test may mean nothing. I was talking overall. By and large they will tell you something.

Now, a more typical case, and this one is close to typical, was a young woman -- this was a man, the first one -- a young woman who had a very good high school record in one of our very demanding schools, had a -- she was from a city in the south -- she had a set of test scores that were so low that when I talked to her about them she said she didn't know whether she would commit suicide. She might as well when she saw those test scores, because her whole life was over. She'd never get in a good college. She

wasn't going to be able to do anything. The totals were somewhere around 900. She subsequently was able, because of her strong academic performance in secondary school, to get into an Ivy League college, to go on to win a prize that they have for graduating seniors, for two years of study in an English university. It's like a Rhodes, but it's not called that. It's a local thing at that university. To come back and go on, to get into a good law school, to finish there with distinction and to pass the bar the first time and again now is clerking for a federal district judge.

KALISH: You're saying that college admissions counsellors should use the SAT scores as just one of a number of criteria. How important or how much, how significant a criteria should it be compared to high school grades and other factors?

BOYD: I think it is one of the most important. I would say if you are going to take -- I'd say the average college probably looks at eight to ten factors. There is no top college in this country that is using the SAT score on this cut off, rigid basis to say, "We line up all the applicants, and in order of their test scores, and that is the order we admit them." That is absolutely false, even though many people are trying to convince the American public that it's true. It is not true.

REINHARDT: William Boyd, President of A Better Chance, interviewed for Options in Education by John Kalish.

MERROW: Boyd may be right. Most college admissions people probably realize that the SAT is coachable, that the score is not precise.

REINHARDT: Most admissions people must know that the Scholastic Aptitude Test is not an aptitude test, but a measure of general achievement.

MERROW: So how to explain what the Admissions Director of an elite privat eastern college told us about the controversy over the use of the SAT.

AN ADMISSIONS DIRECTOR WHO STILL FAVORS STANDARDIZED TESTING

DIRECTOR: I think that it would be a real disservice to many students if the whole idea of a nationally based or nationally normed aptitude test were somehow or other demeaned or discarded. I think it's absolutely vital to have an accurate aptitude test which we can use as only one portion of the admissions process. It's our hope that people would focus on how test scores are used, and not the fact that they may not be accurate or they may not be -- not represent the right kind of judgment.

REINHARDT: Does the test maker bear responsibility for that apparent misunderstanding of the SAT? We asked ETS President William Turnbull about misunderstandings and misuses of the SAT.

WILLIAM TURNBULL, PRESIDENT OF ETS

TURNBULL: It's quite possible to overuse it, to make fine distinctions among students on the basis of small SAT scores. That's a total misuse of the examinations.

MERROW: What do you do to protect against that happening?

TURNBULL: One of the things that we do is say, as I am saying now, that over -- that excessive reliance on the test is bad business, that paying attention to small differences on the test scores is bad business. When I make that comment, I'm not trying to roll with the punches or accomodate to current criticisms of the test. I'm saying yes, and anybody who doesn't believe that has a lot to learn about the fact that human beings are changeable.

MERROW: But you're not doing very much. All you're doing is sort of kind of a half-hearted lecture. Do you do anything to protect against this happening?

TURNBULL: We do stress that in publications that go to students and to their parents. In virtually everything that we write about the test scores, we stress the fact that they are not precise, exact, that they are not infallible, and that other information should be used. We've been saying that since ETS was set up, more than 30 years ago.

REINHARDT: William Turnbull, President of Educational Testing Service.

MERROW: ETS does not prevent colleges from misusing the SAT. It does not require colleges to explain how they use the scores.

REINHARDT: ETS works for the colleges, through the college board. Students pay to take the SAT and the colleges get the information free.

MERROW: SAT scores may be imprecise, coachable and strongly correlated with parental income and social class, but they are from a college's point of view, cheap and convenient.

REINHARDT: Under those circumstances, the question of the SAT's essential fairness seems to get lost.

MERROW: A complete printed transcript of all six parts of this series on standardized testing is available for $6.00.

REINHARDT: The set of cassette recordings costs $18.00.

MERROW: Write to us at National Public Radio - Education, Washington, D. C. 20036.

REINHARDT: $6.00 for the printed transcript of all six programs. $18.00 for the set of cassette recordings.

MERROW: And our address once more, National Public Radio - Education, Washington, D. C. 20036.

REINHARDT: Options in Education is a co-production of National Public Radio and the Institute for Educational Leadership of the George Washington University.

MERROW: Support for the series is provided by the National Institute of Education and the Corporation for Public Broadcasting.

REINHARDT: John Merrow is the Producer of Options in Education. Mary Rowse is our Associate Producer, and Rebecca Goldfield is our Assistant Producer.

MERROW: Barbara Reinhardt serves as co-host and reporter, and we've had engineering assistance from Terry Knight and Marty Kursias.

REINHARDT: Some of the material for this program came from John Kalish in New York City and Pat Bodner in Boston. I'm Barbara Reinhardt.

MERROW: I'm John Merrow.

REINHARDT: This is NPR, National Public Radio.

Program No. 284: Standardized Testing, Part Five

TABLE OF CONTENTS

	PAGE
INTRODUCTION	1
SOUND FROM A MACHINE AT ETS SCORING STANDARDIZED TESTS	1
CAROL KOVACS OF ETS EXPLAINS THE PRECAUTIONS ETS TAKES TO MAKE SURE THE MACHINE MAKES NO MISTAKES	1-2
STEVE SOLOMON OF THE COMMITTE FOR FAIR AN OPEN TESTING EXPLAINS WHY THERE IS A NEED FOR A NATIONAL TEST DISCLOSURE LAW	2-3
STANLEY KAPLAN, HEAD OF THE 90 KAPLAN EDUCATIONAL CENTERS AROUND THE COUNTRY	3-4
STEVE SOLOMON DISAGREES ABOUT WHY PEOPLE AREN'T TAKING ADVANTAGE OF THE LAW	4
WILLIAM TURNBULL, PRESIDENT OF EDUCATIONAL TESTING SERVICE	5-6
LESLEY COLLINS, CHIEF LOBBYIST FOR ETS, SAYS THE GOAL OF TRUTH-IN-TESTING SPONSORS IS TO ABOLISH TESTING	6-7
STEVE SOLOMON DISAGREES	7-8
CONGRESSMAN TED WEISS, (D-NY), PLANS TO RE-INTRODUCE A BILL IN CONGRESS THIS YEAR CALLING FOR FEDERAL LEGISLATION ON TEST DISCLOSURE	8-10
CHUCK STONE, FORMER DIRECTOR OF MINORITY AFFAIRS AT ETS, WANTS A PUBLICLY ACCOUNTABLE ORGANIZATION TO MAKE TESTS	10-11
WILLIAM TURNBULL SAYS ETS WOULD WELCOME AN INVESTIGATION, BUT THAT MOST OF THE CONCERN ABOUT TESTS HAS PEAKED	11-12
ALAN NAIRN, PRINCIPAL INVESTIGATOR AND AUTHOR OF "THE REIGN OF ETS: THE CORPORATION THAT MAKES UP MINDS"	12-13

"Copyright © 1980 National Public Radio and the Institute for Educational Leadership. This material has been furnished for non profit, non broadcast purposes and maynot be duplicated or otherwise used in violation of the Copyright Law."

STANDARDIZED TESTING, PART FIVE

REINHARDT: Every year, millions of students take machine scored, multiple choice, standardized tests, like the SAT, the ACT, the PSAT, the CAT and so on.

MERROW: Ever wonder what happens after you've finished filling in the little boxes with a number two pencil and turned in your answer sheet?

REINHARDT: A machine takes over. This one takes two-fifths of one second to read an answer sheet, 9,000 answer sheets an hour.

SOUND FROM A MACHINE AT ETS SCORING STANDARDIZED TESTS

WORKER: This is the machine that only transcribes what's on the answer sheet onto a magnetic tape. That magnetic tape goes into the computer, where a key is applied, and that is where the actual scoring takes place. But this is transcribing what the candidate put on the sheet, onto a magnetic tape. And they're reading at about 150 answer sheets a minute, is about the speed of that machine, although it can go a little faster than that.

REINHARDT: In most cases, all the test taker ever gets back is a set of numbers. No answer sheet, no analysis of what she did right or wrong.

MERROW: And no way of making sure the machine was accurate. I'm John Merrow.

REINHARDT: I'm Barbara Reinhardt, for NPR's Options in Education, with the fifth part of our report on standardized testing.

MERROW: In this report, the pressure to open up testing. Should test takers have a right to more than a score? Should they be able to compare their answers with the correct ones?

REINHARDT: Should they get a copy of the test itself, so they can look carefully at the questions? The nation's largest test maker, Educational Testing Service, has a quality control system to check up on the machine doing the scoring.

CAROL KOVACS OF ETS EXPLAINS THE PRECAUTIONS ETS TAKES TO MAKE SURE THE MACHINE MAKES NO MISTAKES

KOVACS: We check at least one out of every 400, as many as eight, because the quality control cases are printed one for every fifty documents processed.

MERROW: So the woman working here is actually going over the whole test?

KOVACS: Yes. She is checking each mark, as the candidate put it on the answer sheet against the mark as it was printed by the scanner, and as long as each one checks, then she knows the scanner is functioning properly.

MERROW: Now, here, she's just checked some guy from Denver, Colorado. Suppose she finds that the machine has made a mistake? How does she know that there's not a mistake in the next one and the one right next to it, and so on?

KOVACS: If she finds that the machine has not picked up a mark, then it will be reviewed to see whether it was an improperly gridded mark. If the candidate did not make the mark proper, then that is the end of it. We know that that's the reason why --

MERROW: It was just that guy's mistake.

KOVACS: Correct. If there's any reason to believe that the machine did not function properly, that it should have picked up that mark, then cases surrounding that particular case are all checked, up to 20 or 30 on either side of that case are checked.

REINHARDT: Carol Kovacs of Educational Testing Service in Princeton, New Jersey

MERROW: ETS now provides a copy of the master answer sheet and the student's own answer sheet, to any test taker who requests it.

REINHARDT: For one of its tests, the Preliminary Scholastic Aptitude Test, ETS has begun sending every test taker a copy of the test itself, the correct answers, and his own answer sheet.

MERROW: ETS is responding to pressure for test disclosure. Supporters call it truth in testing; opponents call it throwaway testing.

REINHARDT: Steve Solomon helped push through the first test disclosure law, in New York State.

STEVE SOLOMON OF THE COMMITTEE FOR FAIR AND OPEN TESTING EXPLAINS WHY THERE IS A NEED FOR A NATIONAL TEST DISCLOSURE LAW

SOLOMON: All truth in testing asks for is that students have the right to understand what their test scores mean in complete detail, to understand that a 500, say, on the SAT doesn't pigeonhole you into such a small numerical figure, that a 500 really means somewhere between a 470 and a 530 and it varies any day, and it doesn't really mean that much.

MERROW: Yeah, well, that's easy for you to accomplish. You just accomplished it. You don't need a law.

SOLOMON: Well, you do need a law, because the testing companies weren't doing that. The testing companies weren't giving students detailed information, and most importantly, they weren't letting students see the test questions after they took it.

Many students, when they go through the testing process, come across questions that are ambiguous, questions that obviously have more than one answer. And yet there's nothing they can do about it, because the test doesn't say why did you pick A. It just says what are you going to pick, A or B.

Researchers never can have access to these test questions, to analyze them for possible cultural bias.

MERROW: But there's a reason, because if you give the test away to the kids who took the test, then you have to make up a whole new one.

SOLOMON: Well, current testing procedures actually, in fact, do make up a significant percentage of the tests anyway, because there is such a thing called coaching courses that try to guess what the questions are going to be on a test and then coach students who can afford the courses to do better on the exams.

REINHARDT: Steve Solomon of the Committee for Fair and Open Testing. The New York State law went into effect in January, 1980. It requires disclosure of all standardized tests used for college or graduate school admissions decisions, on request by the test taker.

MERROW: Coaching for tests like the SAT can raise scores, but it costs money, as much as $400.00. That gives kids who can afford it an advantage.

REINHARDT: Disclosure will give everybody the same opportunity.

MERROW: Stanley Kaplan runs the country's largest test coaching school, a $15 million a year business. He supports partial disclosure.

STANLEY KAPLAN, HEAD OF THE 90 KAPLAN EDUCATIONAL CENTERS AROUND THE COUNTRY

KAPLAN: I'm a supporter of truth in testing, except to see the test, because you can have lots and lots of materials that you can learn from. In fact, I agree that there should be one test a year that's released, so therefore they can have a good sample test. But I'm afraid you'll have the same situation with Regents.

If they gave you millions and millions of tests, people will just memorize the answers to the questions and they'll become answer memorizers. In fact, very interesting, with this truth in testing, a very small percentage of people ever ask for the tests. So it's a lesson in frustration. Here we went to all this trouble of releasing these tests, and very few people are interested.

REINHARDT: Stanley Kaplan. His 90 Kaplan Educational Centers coach 50,000 students a year to take standardized tests like the SAT.

MERROW: So far, students in New York State are not rushing to tsee the test. In fact, opponents of disclosure point to the 5 percent request rate as evidence that the law is a mistake.

REINHARDT: Steve Solomon disagrees.

STEVE SOLOMON DISAGREES ABOUT WHY PEOPLE AREN'T TAKING ADVANTAGE OF THE LAW

SOLOMON: That 5 percent figure, you have to understand, is only the first administration of the SAT under a brand new law.

Secondly, the requests for disclosure depend to a great degree on the way in which the truth in testing provisions are publicized to students. The College Board only tells high school students that they can get their questions and answers by an insert in the registration booklet. In one high school in Queens, New York City, these insert forms were in boxes in the basement of the high school, and that was only discovered that they weren't being handed out to students when parents called my office and complained about it, and eventually called the principal's office, and the principal said, "Oh, my God, we weren't letting students know about the law."

MERROW: You're predicting that if in the SAT's, if the high school kids taking that test had an easy way to say yes, a lot more would say yes?

SOLOMON: Overwhelmingly. But even if they still have to go through this difficult process, the percentage is going to rise, as people become more familiar with the law.

REINHARDT: Steve Solomon of the Committee for Fair and Open Testing. He hopes test disclosure will become nationwide through federal law.

MERROW: Only three standardized tests, used for admission to graduate school, are being disclosed nationally.

REINHARDT: On one of those tests, the law school admissions test, 60 percent asked to see the exam, a far cry from the 5 percent

who asked to see the SAT in New York.

MERROW: The discrepancy, says ETS, is explained by the more aggressive and ambitious nature of prospective law students. SAT takers just aren't all that interested.

REINHARDT: On the other hand, all the law students had to do was check off a box. SAT takers had to fill out a separate form and mail it in.

WILLIAM TURNBULL, PRESIDENT OF THE EDUCATIONAL TESTING SERVICE

TURNBULL: I think the legislation passed in New York was undesirable and is having more bad effects than good effects for the students of the state and for the colleges of the state.

REINHARDT: William Turnbull is President of the Educational Testing Service, ETS.

TURNBULL: Who can be against truth in testing? Certainly I'm not. I'm all for truth in testing. If you call it legislation to regulate testing, then we're coming closer to the position I see. The idea of letting students know the content of the tests is something that I fully support, as does everyone at ETS. The idea that students should have full information about the kinds of questions that they'll meet on the test, have a practice test to familiarize themselves with the kind of material they'll see on the test in advance, all that I think is eminently desirable.

MERROW: Disclosing the PSAT, for example, which you now do.

TURNBULL: Yes, and I also think that it's a very good idea to publish copies of the test as given in the past. But I'm not sure that I would say it's necessary to publish every edition of every test given in the past in order to give people an idea of what's in them. That gets you into a different question. In order to do that, you have to be ready, in effect, to scrap every test after one use, and since these are very difficult and expensive tests to produce, that becomes a very wasteful thing. It's like being ready to throw away your dinner dishes after every use. It would be very nice to be able to get more uses from them, rather than consider every test expendable after one exposure.

And so I think the problem in unlimited publication of all test forms after one use is quite a complicated one and, if carried to its extreme, would pose an enormous burden on the students who, after all, pay for the tests in the last analysis.

The very largest programs, and the ones that we are involved with have managed, by very hard and costly work, to work

around the difficulties posed by the legislation. The cost to students has still been very notable, and we have still had to curtail the numbers of test forms that can be given in a year, so much that I must say the students are having to pay more money for poorer service than was true before the law was passed.

REINHARDT: William Turnbull, President of Educational Testing Service. ETS says that test disclosure makes it difficult to maintain the quality of the test, since new questions must be written for each administration of the exam.

MERROW: Different versions of a test have to be equated. They have to be equal in difficulty if score comparisons are going to be meaningful.

REINHARDT: ETS uses one of the six sections of the SAT exam for equating. The scores achieved in this section are not part of an individual's overall SAT score, and the New York law exempts this equating section from disclosure.

MERROW: In New York, the SAT is now given only eight times a year, down from 14 before the legislation went into effect. The test costs $11.00 in New York, $9.25 everywhere else.

REINHARDT: Critics say that ETS didn't have to cut back the number of test administrations or raise the price that much, but did it to punish New York.

MERROW: Today, 20 other states are considering test disclosure legislation, including ETS's home state, New Jersey.

REINHARDT: Two bills have been introduced there. One, modeled after the New York legislation, provides for disclosure of standardized tests used in college or graduate school admissions decisions.

MERROW: The other bill, introduced by a state senator who is also a high school teacher, calls for the disclosure of all standardized tests, from elementary school on up.

REINHARDT: ETS is already fighting the two New Jersey bills. ETS state lobbyist, Lesley Collins.

LESLEY COLLINS, CHIEF LOBBYIST FOR ETS, SAYS THE GOAL OF "TRUTH-IN-TESTING" SPONSORS IS TO ABOLISH TESTING

COLLINS: It looks like a motherhood and apple pie issue until you get into the actual aspects of the bill that are controversial. And this is not easily seen at first glance.

MERROW: What's controversial? What's tricky?

COLLINS: Well, what is not on the surface with the bill is the long term intent of some of the major proponents of this legislation for the abolition of tests. That's what we've been led to believe, and that's what continues to come out from their statements in their conferences and the kinds of things that we hear.

MERROW: Do you somehow feel that if they can beat you here in New Jersey, ETS's home state, that somehow there's the real battleground?

COLLINS: I wouldn't say it's the real battleground, but I think because we're located here we have to be very considerate of the fact that we are an industry in the state, and we do have a position to take.

MERROW: Someone on the side of the truth in testing legislation says that you, ETS, have already hired one of New Jersey's highest priced and highest powered lobbyists to fight for you folks.

COLLINS: Should I use that statement about if they're getting all the chickens in your hen house you should at least get a gun?

ETS is not an active political organization. We don't have packs. Our people are basically people who do work in educational research. We don't know the political environment that well. And so we are looking for advice from someone who's well respected by legislators, and sometimes it helps to know someone who actually knows how to talk to those kinds of people. And, as you well know, ETS has been accused of being, you know, removed from the public sphere. So we don't know that much about it, and we're trying to learn, and this person is, hopefully, helping us learn.

REINHARDT: Leslie Collins, lobbyist for Educational Testing Service in New Jersey.

The earliest the New Jersey Legislature is likely to act on the legislation would be April or May of 1981.

MERROW: Steve Solomon says ETS is deliberately distorting the issue of test disclosure.

STEVE SOLOMON DISAGREES

SOLOMON: ETS has equated -- asking that more information be available to the public, they've equated that to leading towards the destruction of standardized tests. I don't understand how they equate those two issues. Perhaps they have a fear that once test

questions are available for public scrutiny, then people will actually start looking at the test questions carefully and really criticize the tests. It doesn't say that tests should be abolished. How could you even make that decision without having access to vital information over what the test questions are, over the validity studies, reliability studies. You first have to have all that information out in the public domain, and then intelligent discussions can occur.

I want to enable students to have a better understanding of anything that has such a dramatic impact on their life's goals. Clearly, standardized tests do that.

All law schools that are accredited require an ETS test. Thousands of colleges require the SAT.

MERROW: But you're only talking about big kids. What about little kids?

SOLOMON: That's true, and of course at the elementary level, millions of children across the country are classified and tracked by all types of IQ tests, minimum competency testings. Truth in testing must be extended to that level as well, and that's why, out of the 20 states that are now considering truth in testing as the movement is pushing forward, many of them have proposed that truth in testing provisions be extended to the lower levels of education. And we believe that it should not only be extended to lower levels, but also to the licensing tests. ETS doesn't only test students and young children. They test people who want to be stock brokers, who want to be gynecologists. They even have a test for people who want to be CIA agents.

REINHARDT: Steve Solomon of the Committee for Fair and Open Testing. 20 states are considering test disclosure legislation now, and they could each conceivably pass different legislation. That's one argument used to support a national law.

MERROW: A bill was introduced last year in the House of Representatives, and a Congressional committee held eight days of hearings, but it never came up for a vote.

REINHARDT: One of the bill's sponsors, New York Congressman Ted Weiss, intends to reintroduce the bill in the next Congress. He'd like the test makers to support his bill.

CONGRESSMAN TED WEISS (D-NY) PLANS TO REINTRODUCE A BILL IN CONGRESS THIS YEAR CALLING FOR FEDERAL LEGISLATION ON TEST DISCLOSURE

WEISS: I think that the test givers, instead of fighting our efforts, should be supporting efforts to have a single federal piece

of legislation on the books, with all of their might, because even though we can't federally pre-empt the field, I don't think, the likelihood is that if we had a piece of federal legislation, that there would not be the agitation for every state to come along with its own bill.

MERROW: You're saying, in effect, to the testing industry, hey, folks, wise up and support the federal bill, otherwise you're going to be nibbled to death in all 50 states.

WEISS: Precisely so, and what happens is that each state will have a twist of its own on the legislation, although the basic concepts may be the same. And I think it will make the role, the job of the testing companies nigh onto impossible eventually, to try to contend with the requirements in perhaps as many as 50 states across the union.

MERROW: The test makers, at least at the Educational Testing Service in New Jersey, say that's going to increase the cost. It costs a lot of money to make a new test. If we disclose it, we can't use it again. That test has to be secure.

WEISS: That kind of argument has been dismissed out of hand by not just those of us who are promoting legislation, but by other testing companies who have suggested that cost certainly is not one of the valid arguments.

We had, in the course of our hearings, the American Collegiate Testing, ACT people, who came in and who told us that they use new questions on 100 percent of four out of five administrations of those tests. ETS itself acknowledges that in excess of 70 percent of the questions on any particular test are new questions, and it may be even much higher than that.

So that they're developing new questions in any event. I think that this is really a jerry-built argument that they're using, because they think that it's convenient, or that it will help their argument. The fact also is that although they're supposed to be and are created as a non-profit organization, they have very handsome surpluses which, in a profit making institution would be called profits.

MERROW: Are you skeptical of the fact that they say they lost money last year?

WEISS: If they lost money they didn't lose it on the basis of giving tests. And if that's what they're suggesting, I don't believe it. The fact is that over the course of about 30 years they have grown from the classic one room operation to a magnificent two, three, four hundred acre facility on the outskirts of Princeton,

with a visitor center which is a combination conference center, recreation hall, office building, hotel facility. They have centers and facilities across the country, and most of the income comes from people who take the test and who give them their fees, and then they use those monies for all kinds of other purposes. Some of it may be appropriate, and some, I think, quite inappropriate.

MERROW: You've already won some concessions from the test makers. For example, the PSAT, the Preliminary Scholastic Aptitude Test, is now being disclosed. Several of the graduate exams are now available nationwide. If those things continue to happen, does that mean there's no need for federal legislation or even any state legislation?

WEISS: These concessions last usually so long as the pressure is maintained. And even though the people who are currently making those concessions may be genuine in their commitment to making them, if we don't adopt legislation, they or their successors, three or four or five years down the line may decide that really it's unnecessary to do that. And then we have to go through the whole thing all over again.

REINHARDT: New York Congressman Ted Weiss.

MERROW: Some supporters of what's called truth in testing want to open up the testing process and that's all.

REINHARDT: Others want to reduce or do away entirely with machine scored, multiple choice standardized tests.

MERROW: Some of the critics are going after Educational Testing Service, the nation's largest test maker. ETS is chartered as a non-profit institution, and it sees itself as a public servant.

REINHARDT: But Chuck Stone, who worked there for two years, compares ETS to Jack the Ripper, and worse.

CHUCK STONE, FORMER DIRECTOR OF MINORITY AFFAIRS AT ETS

STONE: They're the Godfather of the testing Mafia. ETS is what I call the master of the bait and switch tactic. When you're in school, you got your test back. You take a test, the teacher would give you the test back, and you'd see where you made your mistakes. With ETS you take the SAT, the GRLSAT, you don't know what you did wrong.

MERROW: Would that satisfy you, though, Chuck, if everyone got the tests back with the answer sheet?

STONE: No. I want to get rid of ETS. No, no, I mean I want

to get a publicly accountable agency. ETS is not accountable to the American people. It is a private corporation that makes its decisions irrespective of what the American people want, what the pupils need and what the elected officials expect, and that's what must be done, to make an organization, not ETS, but maybe a public organization, the way we have public institutions, as the main instrument for measuring student potential. And I think that all publicly supported colleges and universities should not be required, or should be prohibited from requiring the SAT in its admissions process.

MERROW: While you're at it, what else do you want?

STONE: Well, I want a Congressional investigation. I want the Congress of the United States or the General Accounting Office to investigate ETS and the other testing agencies around the country, for possible conflict of interest.

MERROW: How so?

STONE: Well, for example, they get government contracts to assess various programs. ETS was involved in assessing a program which used its test. That's a conflict of interests. It's a prima facie conflict of interests. At the end of the year, ETS publishes a report and they always have what they call excess of income over expenses. That's a profit. They don't call it a profit. They just call it excess of income. But it's usually as much as a million dollars. They make a lot of money. ETS is a monolith.

MERROW: Are you saying that's their goal?

STONE: Yeah, I'm saying that the bottom line for ETS is to make money, not to educate children. They're a corporation. They want to stay in business. This is why they resisted the FTC report that said that coaching studies help, because if that's true then you don't need ETS tests. You can study for the test. So naturally that would hurt their business. I think ETS's bottom line is profits, not education.

REINHARDT: Chuck Stone, former Director of Minority Affairs at ETS, now a newspaper man in Philadelphia. Stone wants ETS investigated, but the Federal Trade Commission, the FTC, has already ruled that it has no jurisdiction, because ETS is a non-profit agency. Right now, no Congressional committee is talking about investigating ETS.

MERROW: What's ETS's reaction? President William Turnbull.

WILLIAM TURNBULL SAYS THAT ETS WOULD WELCOME AN INVESTIGATION, BUT THAT MOST OF THE CONCERN ABOUT TESTS HAS PEAKED

TURNBULL: I would welcome an investigation by any constituted

agency. We have absolutely nothing to hide. Of course I suppose I have to come back and ask you on what grounds. I don't look forward to the disruption of an investigation, but I surely would welcome any kind of question, scrutiny, interest in what we're doing.

MERROW: Not long ago, the President of the College Board likened what's going on, really, to a war. Are you at war?

TURNBULL: Oh, I think that these things go in cycles. It became fashionable for a time to put tests under intense scrutiny. My own sense of that is that most of the concern has peaked now. I think many of the questions that were being asked have been answered reasonably well. I think that it will be very healthy if people will retain an interest in what is done in testing, and I would certainly expect that our activities will continue to be scrutinized. I hope we can continue to give good answers.

With a number of people raising questions there is a sustained kind of attempt to see tests diminished, and in that sense I suppose there are periodic battles. Whether there's a long term war or not may strain the analogy.

REINHARDT: William Turnbull, President of Educational Testing Service.

MERROW: There has been one full scale investigation of ETS, by a Ralph Nader research group. Principal investigator Alan Nairn.

ALAN NAIRN, PRINCIPAL INVESTIGATOR AND AUTHOR OF "THE REIGN OF ETS: THE CORPORATION THAT MAKES UP MINDS"

NAIRN: There is a war in the sense that people who are the consumers of tests, the students, the parents, representatives of groups like minority organizations whose constituents have always been hurt by these tests, are finally saying, you know, "We've had enough. We're now going to make this a political issue and we're going to call ETS into account." There's definitely a major political conflict going on.

The -- ETS has, I think, distorted the nature of that conflict by saying it's an attack on standards and merit itself. Like that's what Hanford said in that speech. He said, "This is a war on the idea of merit and a war on the very idea that we should test people."

MERROW: Is it?

NAIRN: No, not at all. It's very pretentious of ETS to say that their product is the standard of merit. We're not criticizing the assessment of merit. We're saying ETS does not assess merit.

We're saying it does not really test what people can do and that, in claiming that, they're making a false claim.

I mean it's not an attack on the idea of finding out what people know, because that's important in helping people learn, obviously. It's an attack on ETS's false claims that they can measure people's aptitude. Truth in testing is not sufficient to deal with problems of testing by any means. What it does is it provides a tool. It ends the secrecy. It puts the facts out in the public domain, and you can then use those facts to start to make changes in the test system. But it's only the first step, but it's an essential first step.

You see, because if you don't have that, anyone who criticizes the status quo is put in the position of having less information, because the test companies can then turn around and say, "Look, we know what's on these tests. We have these studies. We've seen them and you haven't." Kind of like the Pentagon said during Vietnam. "We have accessed information we can't disclose which shows you're wrong." And what can you say to that? You're put in the position of being systematically less informed than they are, and you've got to even that up, and that's what truth in testing is needed for.

REINHARDT: Alan Nairn, principal author of, "The Reign of ETS: The Corporation That Makes up Minds." The book is the product of a six year investigation of Educational Testing Service.

MERROW: Are there alternatives to standardized testing? That's next in this series. Hope you'll be with us.

REINHARDT: A complete printed transcript of all six parts of this series is available for $6.00. The set of cassette recordings costs $18.00.

MERROW: Write to us at National Public Radio - Education, Washington, D. C. 20036.

REINHARDT: $6.00 for th printed transcript of all six programs. $18.00 for the set of cassette recordings.

MERROW: And our address again, National Public Radio - Education, Washington, D. C. 20036.

REINHARDT: Options in Education is a co-production of National Public Radio and the Institute for Educational Leadership of the George Washington University.

MERROW: Support for the series is provided by the National Institute of Education and the Corporation for Public Broadcasting.

REINHARDT: John Merrow is the Producer of Options in Education. Mary Rowse is our Associate Producer and Rebecca Goldfield is our Assistant Producer.

MERROW: Barbara Reinhardt serves as co-host and reporter, and we've had engineering assistance from Greg Hedgepath and Dave Glasser.

REINHARDT: I'm Barbara Reinhardt.

MERROW: I'm John Merrow. This is NPR, National Public Radio.

PROGRAM NO. 285: STANDARDIZED TESTING, PART SIX

TABLE OF CONTENTS

	PAGE
INTRODUCTION	1
A STUDENT TEACHER TELLS ABOUT A STUDENT WHO DREADED THE TEST	1
JUDAH SCHWARTZ FINDS THE TESTS USELESS	1-2
A PRINCIPAL WHO HAS STUDENTS READ TO HIM AS A MEANS OF KEEPING TRACK OF THEIR ABILITY	3
EVA BAKER THINKS TESTING OUGHT TO BE A LOCAL MATTER AND NOT LEFT TO THE "EASTERN TEST GURUS"	2-3
HOWARD WAINER SAYS COMPUTER TECHNOLOGY CAN MAKE TESTING A MORE EXACT AND HIGHLY PERSONALIZED PROCEDURE	3-4
FOR BANESH HOFFMAN, THE PROBLEM STEMS FROM AN "IDIOTIC COMPULSION" TO RATE PEOPLE NUMERICALLY	5-6
JUDAH SCHWARTZ SAYS THE QUEST FOR THE RIGHT ANSWER LEADS TO INFERIOR EDUCATION	6-7
STUDENTS READING ALOUD TO TEACHER JUDY BUCHANAN	7
BUCHANAN DESCRIBES HOW SUCH READING HELPS HER EVALUATE THE BOYS' SKILLS	7-8
OTHER TEACHERS SUGGEST METHODS OF WORKING WITH BUCHANAN'S STUDENTS AT A STAFF REVIEW SESSION	8-9
JUDY BUCHANAN ON WHY SHE PREFERS THIS METHOD TO TESTING	9-10

"Copyright © 1980 National Public Radio and the Institute for Educational Leadership. This material has been furnished for non profit, non broadcast purposes and maynot be duplicated or otherwise used in violation of the Copyright Law."

STANDARDIZED TESTING, PART SIX

A STUDENT TEACHER TELLS ABOUT A STUDENT WHO DREADED THE TEST

TEACHER: I had a student who didn't want to take the standardized test. He had had trouble learning to read and was finally making real progress, but he was smart enough to know that he wasn't going to be able to do that test, and he didn't want to sit there and not be able to do it. And in the end he did it. I said he could have lunch with me on the day that he finished the test, and he did it and he had lunch with me. But it was a real struggle for him, and he already knew what he knew, and he knew what he didn't know, and I'm sure that his parents know that too, and there was just no need for this. He didn't need to be compared to a child in the rest of the country.

MERROW: That student's discomfort is only one reason some folks would like to get rid of standardized tests, those nationally normed, multiple choice, machine scored exams that compare students with others in the same grade throughout the country.

REINHARDT: Do the problems lie in the tests themselves, or the way tests and test scores are used?

MERROW: Or do the problems lie deeper, in our own need to label each other with a number, to rank each other according to ability?

JUDAH SCHWARTZ FINDS THE TESTS USELESS

SCHWARTZ: I think the only sensible thing is to compare the child to the task and not the child to another child. When you take your child to a physician, it is of little interest to you to know that this child has 73 percent of the number of bacteria that the average child in town has. You want to know whether or not that child has an infection, and if that child has an infection you want that child to be treated in a way that is idiosyncratically appropriate.

Let me give you my analogy. Suppose that psychometricians were the people who ran blood tests. Well, what you would do is you would go and get a blood test and then they would take the numbers for white blood count and red blood count and sedimentation rate and hemoglobin and triglycerides and lipids, and they'd average them all together and you'd get a score on a blood test. And it's about as useful as that.

REINHARDT: Judah Schwartz, a professor at the Massachusetts Institute of Technology. I'm Barbara Reinhardt.

MERROW: I'm John Merrow. for NPR's Options in Education, with the final part of our series on standardized tests in the schools.

REINHARDT: In this report, a search for alternatives to the way we use them now.

A PRINCIPAL WHO HAS STUDENTS READ TO HIM AS A MEANS OF KEEPING TRACK OF THEIR ABILITY

STUDENT: "He's still out there making awful faces, the woman cried angrily. Now he is sticking out his tongue at us. He certainly is not a polite boy."

PRINCIPAL: All right, Jackie. Why do you think they call this story, "Funny Faces?"

STUDENT: Because he was making funny faces in the mirror.

PRINCIPAL: And why was he doing that?

REINHARDT: William Corbett is Principal of Lowell Elementary School in Watertown, Massachusetts. Twice a year he has each of his 300 or so 1st through 4th graders read to him.

MERROW: Corbett says his contact with the children tells him more about their reading ability than the standardized tests alone ever could.

REINHARDT: Eva Baker took the idea a step further when she was testing the reading performance of children in California. She had them read aloud and kept careful track of the kinds of errors they made.

EVA BAKER THINKS TESTING OUGHT TO BE A LOCAL MATTER AND NOT LEFT TO THE "EASTERN TEST GURUS"

BAKER: Lo and behold, the reading performance of California primary age children looked very good, as opposed to the bad news that one normally gets on the standardized tests, which is in part an artifact of those tests, because half of the group always has to be below average.

So the question, is there a better way, I think, also has to be asked, and it has to be asked without the sense that there are so-called gurus someplace, probably on the East Coast, really know how to write test items and really understand this process at a very sophisticated level. I'd like personally to see a lot more activism and a lot more local development of measures, because I think that the benefit is having people buy into the process and own it, and therefore try and do direct instruction in a much more consistent way, than if it's a test that's made elsewhere, comes in

once and is gone.

MERROW: Eva Baker of UCLA.

REINHARDT: Standardized testing has expanded in the last decade or so, and the major reason is the improvement of computer technology. This year, our 50 million students will take somewhere around 200 million standardized tests.

MERROW: Computer experts say that the same technology that made testing the masses possible can now be used for finely tuned testing of individuals.

REINHARDT: Most paper and pencil standardized tests are given to a broad range of students, so the range of questions has to be broad, too, from very easy to very hard.

HOWARD WAINER SAYS COMPUTER TECHNOLOGY CAN MAKE TESTING
A MORE EXACT AND HIGHLY PERSONALIZED PROCEDURE

MERROW: What if you could set the student down in front of a computer terminal? Howard Wainer.

WAINER: The computer would say, hi, what's your name? And you'd write down your name, and then it would ask you a question sort of in the middle of the ability range. And you would try and answer it. If you could answer it easily, then the computer would then pick another question from that pool that would be more difficult, and see if you could answer that one. If you could, it would pick a much more difficult one still, and keep doing that until you started to miss. Then it would back up a little bit.

And that way -- and keep doing that until it zeroed in on your ability. And that way, you would never be bothered, or would be bothered for a really minimal amount of time, with questions that were really trivially easy. And so you wouldn't have time to get bored or fatigued. Also, you wouldn't be struck with a whole lot of absolutely impossible questions that you had no chance of answering. If your knowledge of French was just bonjour, ou les toilettes, or le vesees, it would never say to you something in very complex form, and ask you to understand the subjunctive.

And so you would always -- once you got into the test, and it would happen rather quickly, you would always be operating right at your ability. So the questions would always be difficult, but never impossible. There would be a challenge all the way through it, and it could come out with a much more accurate estimate of your ability in a more limited amount of time.

GOLDFIELD: In what way is this an improvement over a good teacher who could just watch my progress through a year, take notes, keep files, have conferences?

WAINER: Well, it's a lot cheaper than a good teacher, and a lot better than a poor teacher. The principal advantage is that it will behave like a good teacher, but it will behave like a good teacher in a consistent way.

And so, if it's being used for any sort of standardized evaluation, we'll have a good measure of being able to

make comparisons across individuals, and also be able to get a good estimate of each individual.

It also has enormous capability for going in and helping the student as a learning aid, because just as it can ask questions, it can also answer them. And if you make a mistake, it can tell you what the correct answer was. It can act just like a good teacher.

GOLDFIELD: Suppose I took the test. How would the computer interact with me?

WAINER: Well, this is all hypothetical. But my guess is, you could be asked, would you prefer a man or a woman? Let's think of this as something in quite the distant future, in which the computer is very much like Hal in 2001.

We're still talking about computerized adaptive testing, but what I'm going to do is, I'm going to add on one development which is not yet available. Right now, computers can speak and they can understand when you type in responses. But let's assume for a moment that you can speak into them and they'll understand you. That's far away right now. Or you would just type it in.

GOLDFIELD: How far away is that, anyway?

WAINER: I don't know. Let's say, you go in and the computer says to you, would you prefer to speak with a man or a woman? And you can make your choice. You might feel more comfortable with a man or you might feel more comfortable with a woman. And then the voice of the computer would then change, depending on what you would prefer.

Of course, if it's as test of reading comprehension, the computer wouldn't be speaking to you, it would be writing things out, and you'd be reading them. If it was a test of mathematical concepts, it might be using some sort of screen in which pictures and graphs would be drawn.

In general, it would be using all sorts of facilities. It can present things in very complex ways, it can present things moving. There's a project that we're doing at Harvard now, with Peter Euber, in which we're using a computer to present things in a dynamic way, so that the displays are constantly moving in front of you.

So the computer as a mode of presentation is like a slide show and a stereo all rolled up in one, with enormous amounts of information, and present things in enormously interesting ways. And as it presents this stuff to you, you give responses. It can be programmed to help you out, to --

GOLDFIELD: How would that work? Let's say I'm suddenly just feeling really flustered and things are going poorly? Could I make that known to my computer friend, here?

WAINER: Some computers, when you make a mistake, say you made a mistake, but don't worry, I understand what you mean. But it can be programmed in any way. It can say, okay, listen, that question might have been too rough for you, let's try something a little bit

easier, or something that's more up your line. Or you can have your computer to say any number of things to be sure to keep you -- to make you feel as comfortable as possible. Because, remember, the idea of testing is not to see how you perform under stress, rather, to see what you know and how you know it.

REINHARDT: Howard Wainer, of Educational Testing Service, talking with Rebecca Goldfield.

FOR BANNESH HOFFMAN, THE PROBLEM STEMS FROM AN "IDIOTIC COMPULSION" TO RATE PEOPLE NUMERICALLY

MERROW: Wainer has tremendous faith in the tests and the ability of computers to improve them. Bannesh Hoffman has faith in neither.

HOFFMAN: Essentially, it's that's neurotic compulsion to get a number out of something that drives people back from the subject to the purely mechanical multiple choice form. Why should you, if you've found -- let me ask you this: Take Kafka and take Shakespeare and take Arnold Bennett, and would you please rate them numerically for me?

MERROW: I'd give you an 87.3 and -- I understand.

HOFFMAN: Do you know, there was an attempt to grade essays by computer? You know about it?

MERROW: Tell me.

HOFFMAN. Well, the people -- and, incidentally, the College Board was sufficiently impressed with the prospect that it gave them handsome amounts of money to continue their research. And they had people write essays, and they had essay graders grade the essays. And then they put on the computer all the things that they could think of about the various essays and the grades that they got.

And the computer came out with a whole set of criteria for grading essays. I think there were 31 different criteria. And I remember the top five, I think I can recall them. One was the length of the essay. That - in words, I presume. And another was the number of words on a common word list. If you had too many of these usually used words, that was counted against you. Then there was the number of commas, the average length of the words, and they put this into a computer again, and this time the computer graded essays and they pointed out, see, you can't tell which essays were graded by computer -- what scores were given by computer and what scores were given by the human readers.

And for this statistical reason, they thought they had found a way of evaluating essays without using human graders, and doing it at least as well. One interesting thing is that their computer gave grades to two decimel places, from zero to five or one to five, I've forgotten which, so that you could have a grade of 3.29.

Now, that means that you are grading essays on a scale with 500 or so different levels. And no human being, no enormous committee of human beings would ever come up with such a

refined sort of rating.
MERROW: But if I got that grade, and I was up to be considered for something, they would look at that and they'd say, 3.9? This guy only got a 3.27, let's take Merrow.
HOFFMAN: The second decimal place would be the crucial one.
REINHARDT: Bannesh Hoffman, author of The Tyranny of Testing.

JUDAH SCHWARTZ SAYS THE QUEST FOR THE "RIGHT ANSWER" LEADS TO INFERIOR EDUCATION

MERROW: When you take a standardized test, you normally stop your classroom work and spend an hour or so answering lost lists of questions.
REINHARDT: Your test taking is separate from your learning. According to Judah Schwartz, that's bad. For him, a test question ought to go something like this.
SCHWARTZ: I pose to you the following problem. Estimate the number of sprinkles on a chocolate cone. Any child who can evolve a reasonable strategy for approaching that problem, I don't care about the answer, but a strategy for approaching that problem, is a child that has mastered something that is structurally very important and very deep about dealing with the quantitative aspects of the world in which he or she lives.
ROUSE: Okay, so you have identified the fact that the strategy, the thought process is far more important than the final result. But doesn't there have to be some way of getting a right answer, after -- assuming a strategy is accurate?
SCHWARTZ: Right answer is a funny notion, because we ordinarily think of mathematics as being the world of right answers. Let me ask you the following problem, that comes from the daughter of a colleague of mine.

There are three people in our family who like hard boiled eggs, and there are three people in our family who like soft boiled eggs. How many people are there in my family? Now, you will recognize instantly that that question doesn't have an answer in the normal sense of the word. It does, however, have a quantitative answer. The answer is at least three, but possibly 73.

Now, I think it is terribly important that one get away from the notion of "the right answer" as part of what mathematics is about.
ROUSE: Could you give me more examples of what you would tell a teacher, or what you would do as a teacher, to assess a child's progress and knowledge in mathematics?
SCHWARTZ: Let me give you an example of something that I do with my students at MIT, but which I see no reason in the world for not doing with a fourth grade class.

Suppose that our subject in fourth grade mathematics was the addition of decimals. I would ask the children to make up a quiz on the addition of decimals, and defend why they felt that that quiz would be a good quiz. That forces them to stand back a bit and

reflect what is it about this material that is important for us to know, and how can we know it?

And the reason that's an important question to ask is because in the long run, the aim of our formal educational system is to get people to the point where they can educate themselves. Unless we give them practice in asking the question what is it important for me to know, and how do I know if I know it, they won't learn that.

ROUSE: What would your ideal tests be like if you were a fourth grade teacher?

SCHWARTZ: They would be indistinguishable from instruction. Every good test teaches the child something, and every good instructional interaction provides a measure of assessment of how that interaction is going.

REINHARDT: Judah Schwartz, a professor at MIT, talking with Mary Rouse.

MERROW: Alternatives exist in practice as well as in theory. In one Philadelphia public school, parents can choose to place their children in classrooms which strongly de-emphasize testing.

STUDENTS READING ALOUD TO JUDY BUCHANAN

REINHARDT: Except for one required standardized test a year, and occasional basic skills quizzes, children in these classrooms are free of traditional testing.

MERROW: Instead, teachers rely on interviews with the children and direct observation of their skills.

CHILD: Patty and Fay were playing. Look at the party and the --

BUCHANAN: Pretty.

CHILD: Pretty pin. I found it, said Patty. It is a pretty pin, said Fay. See the way to look in the sun.

BUCHANAN: Good reading.

CHILD: I like that pin, said Fay. It looks like a good pin.

BUCHANAN DESCRIBES HOW SUCH READING HELPS HER EVALUATE THE BOYS' SKILLS

REINHARDT: Can you tell me, what did you learn today as you were listening to those two boys read? What did you learn about the first one?

BUCHANAN: Okay, I already knew part of this before. He's coming along as a reader. He makes very few errors. When he does make a mistake, he corrects it and I help him . He corrects it, and I help him correct it. It think he's had the name -- the name of the girl was Patty, and the word describing the pin was pretty, and he got confused. And I told him it was pretty, and I heard him use it correctly throughout that page that he read.

So he understood the correction and he was able to use it properly after that. He didn't make mistakes on sight words, like when, there, this, that. He seems to know those pretty well. He read with expression, he seemed to know what he was doing. That tells me, not just because I heard him today, but because I've heard him since September, that things are going along well for him. He's

grasping what he needs to be grasping, he's making sense of it. And he even seems to be enjoying it still, which is good.

REINHARDT: What did you learn about the second boy?

BUCHANAN: Again, I know something about him because I read with him since September, too. He does have trouble with those sight words. And he often makes up a different one to put in. I didn't find this out today because we really didn't have time. But he still somehow understands what he's read, even though he doesn't get the sight words right. He gets the words like pretty pin and Grandpa and do you want to play a game. He understands what's happened in the story, even though he's made a lot of errors. Somehow he's still getting the meaning from the words, which is the most important thing.

He also gets the first part of words right. For bet, he said better. He hasn't probably read the word bet too often in the story. He's read the word better, though. And he made a mistake, but it was a mistake that had meaning and was connected. He just added an extra syllable. But he's still processing the words, he's still doing a lot. He just hasn't quite got that word down yet.

REINHARDT: I'm trying to think if you could have learned that on a standardized test.

BUCHANAN: Not on any that I've seen. I wouldn't have learned that. Because he would have either had to choose the right answer from a group of answers, which would have been different from thinking in his own head, or he would have had to choose the word that was not correct, or he would have had to right something down. And none of those are exactly what he had to do when he read that. So I don't know that I would have learned that about him.

But even if I could have learned it about him, it wouldn't have been any less time than I learned it this way. And it probably would have been more time between giving the test and waiting for results, or whatever. It really didn't take me long to learn that today.

OTHER TEACHERS SUGGEST METHODS OF WORKING WITH BUCHANAN'S STUDENTS AT A STAFF REVIEW SESSIO

MERROW: Judy Buchanan, a fourth grade teacher at the Lee School in Philadelphia. She assesses her children's progress by observing them. This alternative to traditional tests includes weekly staff meetings, at which teachers discuss their students and share suggestions.

BUCHANAN: I think that he can get meaning from stories, but since he misses some sight words that are crucial to meaning, when you get beyond the I Can Read book stage, those sight words carry meaning, too, like, if he doesn't understand that a question is being asked in a story because he doesn't know what or who or will, and he says well for it, or something like that, he may miss the meaning.

TEACHER: Well, I think that those words are structural words, not -- that they do have meaning, but that their meaning lies in the structure in which they're used. And I notice that in his writing, he doesn't use them, either. He -- the only connectives he has in his writing are then and and, and a little bit of but.

And maybe he needs to practice saying them, making up sentences. There could be gains about -- I keep trying to think of -- what's that one about, who did the something? Who stole the -- that kind of thing. I don't know how familiar he is with those words, or with the structures. From his writing, it seems like he maybe isn't. In the book, this Clifford book which you said he liked, there's not a one of them.

So, I don't know, maybe some of that. But I would suggest not the words by themselves, but in sentences. What happened when this, who did that, and maybe he could make those up, too.

TEACHER: I'm not sure if this sort of thing is available, but it seems to me there must be books or games or maybe you could make picture cards with words on the bottom, where the meaning lies in the little words, not in the big ones.

For instance, a large picture of a field full of birds, and words -- pick a bird, any bird, a bird with "a" emphasized. And then another picture full of cows and snakes and bears and things like that, but only one bird, the bird, the bird. That sort of thing.

What gave me the idea is Jack Hemp's Hop Skip Jump book. That book is on the words that he doesn't have trouble with, but maybe that idea could be used.

REINHARDT: This kind of assessment, how could a school district use this to show whether or not it's doing all right?

BUCHANAN: Well, I think it requires a great deal of trust. I think -- and I think that's why we don't use it so much. It requires that a teacher be believed and that the parents trust themselves in their own judgments of their children.

REINHARDT: But you've mentioned parents and children. And, of course, they're not the people who make the decisions, nor are they the people always whom the decision makers have to please.

BUCHANAN: That's true, they don't have the power in the system. I'm just not sure why anymore we want to know what we can find out from standardized tests. I know what my children can do. I don't need to know that they can't do it as well as some other child. It doesn't help me and it doesn't help them.

And I don't think that it helps anyone. But because it's a number and because it is "standardized", people can believe that it has meaning, where if I say to someone, well, this is what your child did and this is how your child did it, and this is how well your child is doing and what your child needs to work on more, there's no way to standardize that. And I think that in a small place, with people holding on to records from year to year, with people working together, with teachers talking to teachers from the

previous year, and with parents being able to talk with the teachers and observe their children and look at their children's work, that this kind of assessment can really work, because it helps you teach the child what the child needs to learn.

It does not tell you how your child is doing compared to some other child in some faraway place. And I don't think we need to know that.

MERROW: Judy Buchanan, a fourth grade teacher at Lee School in Philadelpia.

REINHARDT: Doing away with nationally normed standardized tests would put heavier demands on teachers. They'd have more responsibility for individually assessing students. Are they prepared for that responsibility?

MERROW: Parents and school boards are accustomed to numbers and percentile rankings. Would they -- we -- be able to kick the habit?

REINHARDT: But perhaps we shouldn't be too quick to say standardized tests are worthless. Their scores do give an idea of how the children in one district compare with those in another, on certain skills.

MERROW: Nationally-normed standardized test results help college admissions officers determine whether an A from one high school is as good as an A from another.

REINHARDT: But remember, we can get that information by giving the nationally-normed test to only a small sample, less than ten percent of students in a school district. If the sample is selected randomly, the results apply to the whole district.

MERROW: And there won't be individual scores and individual percentile rankings to misuse afterwards.

REINHARDT: A complete printed transcript of all six parts of this series is available for six dollars. A set of cassette recordings costs $18.00.

MERROW: Write to us at National Public Radio - Education, Washington, D.C. 20036.

REINHARDT: Six dollars for the printed transcript of all six programs, $18.00 for the set of cassette recordings.

MERROW: Our address once more - National Public Radio - Education, Washington, D.C. 20036.

REINHARDT: Options in Education is a co-production of National Public Radio and the Institute for Educational Leadership of the George Washington University.

MERROW: Support for the series is provided by the National Institute of Education and the Corporation for Public Broadcasting.

REINHARDT: John Merrow is the producer of Options in Education, Mary Rowse is our associate producer, and Rebecca Goldfield is our assistant producer.

MERROW: Barbara Reinhardt is our co-host and reporter, and we've had engineering assistance from Mary Ishnaga and Greg Hedgepath.

REINHARDT: Some of the material for this program came from Liz Reis

of WBUR Boston, and Lisa Schlien in Los Angeles.
I'm Barbara Reinhardt.
MERROW: And I'm John Merrow.
REINHARDT: This is NPR -- National Public Radio.
(End of proceedings as recorded.)